TRANSACTIONS

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

VOLUME 43

JANUARY TO JUNE, 1938

PUBLISHED BY THE SOCIETY
MENASHA, WIS., AND NEW YORK
1938

Chester C. Corbin Library Fund
Bound July, 1938

Composed, Printed and Bound by
The Collegiate Press
George Banta Publishing Company
Menasha, Wisconsin

TABLE OF CONTENTS

VOLUME 43, JANUARY TO JUNE, 1938

POSTULATES FOR ABELIAN GROUPS AND FIELDS IN TERMS OF NON-ASSOCIATIVE OPERATIONS*

BY

B. A. BERNSTEIN

1. **Object.** The object of this paper is to present sets of postulates for abelian groups and fields in terms of the non-associative (and non-commutative) operations "$-$" and "/", the inverses of "$+$" and "$\times$" in a field. The postulates will thus treat directly the properties of the inverse operations in a field, properties important from the standpoint of operations in general, but perhaps not sufficiently emphasized in the usual treatment of groups and fields. Three sets of postulates will be given for fields. In each set, the postulates free from "/", taken by themselves, will form a set of postulates for abelian groups. Unlike other sets of postulates for fields known to me, the sets offered contain no (unconditioned) existence proposition other than one demanding that the class contain at least two elements. The consistency, necessariness, and sufficiency of the postulates are established by the usual methods. The postulates will be found to be simple and "natural".

2. **Postulates (F) for fields.** A *class* K of elements will be a (non-trivial) *field* with respect to a pair of *binary operations* "$-$", "/" if K, $-$, / satisfy the postulates N, S_1–S_3, D_1–D_6 following:

N. *K contains at least two distinct elements.*

S_1. $$a = a - (b - b),$$
if a, b, and the combinations indicated are in K.

S_2. $$a - (b - c) = c - (b - a),$$
if a, b, c, and the combinations indicated are in K.

S_3. *If a, b are in K then $a-b$ is in K.*

D_1. $$a = a/(b/b),$$
if a, b, and the combinations indicated are in K.

D_2. $$a/(b/c) = c/(b/a),$$
if a, b, c, and the combinations indicated are in K.

* Presented to the Society, April 11, 1936; received by the editors February 23, 1937.

At the time of the reading of the paper no other postulate set for groups or for fields in terms of "indirect" operations was known to me, except Wiener's postulates for fields in terms of "@" (Publications of the Massachusetts Institute of Technology, June, 1920). Since then, there have come to my attention papers on groups treated from my general point of view by R. Baer and F. Levi (Zentralblatt, vol. 4, p. 338), Morgan Ward (these Transactions, vol. 32, p. 526), G. Y. Rainich, D. G. Rabinow. Rainich's paper and Rabinow's are, I understand, in press. I have modified the original draft of my paper so as to put in the background possible duplications of the results of these writers.

D_3. *If a, b, $b-b$ are in K, and if $b \neq b-b$, then a/b is in K.*

D_4. *If a, b, $b-b$ are in K, and if $b=b-b$, then a/b is not in K.*

D_5. *If a, b, and the combinations indicated are in K, and if $a/b=a/b-a/b$, then $a=a-a$.*

D_6. $$(a-b)/c = a/c - b/c,$$
if a, b, c, and the combinations indicated are in K.

3. **Consistency, necessariness, and independence of Postulates (F).** The following arithmetic system $(K, -, /)$ satisfies all the Postulates (F):

$$K = 0, 1;\ a - b = a + b \pmod 2;\ a/b = a \div b \pmod 2.$$

Hence, *Postulates* (F) *are consistent.*

The properties given by Postulates (F) are seen to be properties of "subtraction" and "division" in a field. Hence, *Postulates* (F) *are necessary for a field.*

Finally, *Postulates* (F) *are mutually independent.* The independence-systems are given by the table below. The systems in this table are all arithmetic. Some of the systems are modular, the modulus being enclosed in parentheses following the operations. A system contradicting a postulate P is lettered $\overline{\mathrm{P}}$. The independence-table follows.

INDEPENDENCE-SYSTEMS FOR POSTULATES (F)

System	K	$a-b$		a/b	
$\overline{\mathrm{N}}$	Null class				
$\overline{\mathrm{S}}_1$	0, 1	0		$a \div b$	(2)
$\overline{\mathrm{S}}_2$	0, 1	a		$0 \div 0$	
$\overline{\mathrm{S}}_3$	0, 1	$0 \div 0$		$a+b$	(2)
$\overline{\mathrm{D}}_1$	0, 1, 2, 3, 4	$a+4b$	(5)	$ab+(0\div b)$	(5)
$\overline{\mathrm{D}}_2$	0, 1, 2	$a+2b$	(3)	$a+(0\div b)$	
$\overline{\mathrm{D}}_3$	0, 1	$a+b$	(2)	$0 \div 0$	
$\overline{\mathrm{D}}_4$	0, 1	$a+b$	(2)	$a+[0\div(ab+a+1)]$	(2)
$\overline{\mathrm{D}}_5$	0, 1	$a+b$	(2)	$0+(0\div b)$	
$\overline{\mathrm{D}}_6$	0, 1, 2	$a+2b$	(3)	$ab+(1+2a^2)(b+2b^2)$ $+(0\div b)$	(3)

4. **Theorems.** The theorems T_1–T_{22} following are derivable from Postulates (F). They will establish the sufficiency of Postulates (F) for fields.

T_1. $a=b-(b-a)$.

T_2. $a-b=(c-c)-(b-a)$.

T_3. $(a-b)-c=(a-c)-b$.

T_4. $a-a=b-b$.

DEFINITION 1. $0=a-a$.

DEFINITION 2. $a'=0-a$.
DEFINITION 3. $a+b=a-b'$.

T_5. *If a, b are in K then $a+b$ is in K.*
T_6. $a+b=b+a$.
T_7. $a+(b+c)=(a+b)+c$.
T_8. *For any two elements a, b in K there is an x in K, namely, $x=b-a$, such that $a+x=b$.*

In Theorems T_9–T_{16} and in Definitions 4, 5, 6 following, unless told otherwise, *it is supposed that the elements indicated are all in K*, i.e. (D_3, D_4), it is supposed that no "divisor" is 0. In the proofs of the theorems, "Hypothesis" will refer to this supposition.

T_9. *If $a=0$ then $a/b=0$, and conversely.*
T_{10}. $a=b/(b/a)$.
T_{11}. $a/b=(c/c)/(b/a)$.
T_{12}. $(a/b)/c=(a/c)/b$.
T_{13}. $a/a=b/b$.

DEFINITION 4. $1=a/a$.
DEFINITION 5. $a_1=1/a$.
DEFINITION 6. $ab=a/b_1(b\neq 0)$; $a0=0$.

T_{14}. $1\neq 0$.
T_{15}. $a_1\neq 0$.
T_{16}. $(a/b)_1=b/a$.
T_{17}. $0a=0$.
T_{18}. *If a, b are in K then ab is in K.*
T_{19}. $ab=ba$.
T_{20}. $a(bc)=(ab)c$.
T_{21}. *For any two elements $a\neq 0$, $b\neq 0$, there is an x, namely $x=b/a$, such that $ax=b$.*
T_{22}. $a(b+c)=ab+ac$.

5. **Proofs of the theorems.** The proofs of Theorems T_1–T_{22} follow. In T_9–T_{16}, if $a\neq 0$, $b\neq 0$, $c\neq 0$, then any combination of a, b, c is in K, by S_3, D_3. This fact will generally be understood in the proofs.

Proof of T_1. $a=a-(b-b)=b-(b-a)$, by S_1, S_2.
Proof of T_2. $a-b=a-[b-(c-c)]=(c-c)-(b-a)$, by S_1, S_2.
Proof of T_3. $(a-b)-c=(c-c)-[c-(a-b)]=(c-c)-[b-(a-c)]$ $=(a-c)-b$, by T_2, S_2, T_2.
Proof of T_4. $a-a=b-[b-(a-a)]=b-b$, by T_1, S_1.

Proof of T_5. By Definition 3, Definition 2, Definition 1, S_3.

Proof of T_6. $a+b=a-b'=a-(0-b)=b-(0-a)=b-a'=b+a$, by Definition 3, Definition 2, S_2, Definition 2, Definition 3.

Proof of T_7. $a+(b+c)=(b+c)+a=(b-c')-a'=(b-a')-c'=(b+a)+c$ $=(a+b)+c$, by T_6, Definition 3, T_3, Definition 3, T_6.

Proof of T_8. $a+(b-a)=a-(b-a)'=a-[0-(b-a)]=a-[(c-c)$ $-(b-a)]=a-(a-b)=b$, by Definition 3, Definition 2, Definition 1, T_2, T_1.

Proof of T_9. If $a=0$, then $a/b=(b-b)/b=b/b-b/b=0$, by Definition 1, D_6, Definition 1. If $a/b=0$, then $a/b=a/b-a/b$, by Definition 1; hence $a=a-a=0$, by D_5, Definition 1.

Proof of T_{10}. $a\neq 0$, $b/a\neq 0$, by Hypothesis. Hence $b\neq 0$, by T_9. Hence, $a=a/(b/b)=b/(b/a)$, by D_1, D_2.

Proof of T_{11}. $a\neq 0$, $b\neq 0$, $c\neq 0$, by Hypothesis. Hence, $a/b=a/[b/(c/c)]$ $=(c/c)/(b/a)$, by D_1, D_2.

Proof of T_{12}. $b\neq 0$, $c\neq 0$, by Hypothesis. If $a\neq 0$, then $(a/b)/c$ $=(c/c)/[c/(a/b)]=(c/c)/[b/(a/c)]=(a/c)/b$, by T_{11}, D_2, T_{11}. If $a=0$, then $(a/b)/c=0=(a/c)/b$, by T_9, T_9.

Proof of T_{13}. $a\neq 0$, $b\neq 0$, by Hypothesis. Hence $a/a=b/[b/(a/a)]=b/b$, by T_{10}, D_1.

Proof of T_{14}. $a\neq 0$, by Hypothesis. Hence, $1=a/a\neq 0$, by Definition 4, T_9.

Proof of T_{15}. $a\neq 0$, $1\neq 0$, by Hypothesis, T_{14}. Hence, $a_1=1/a\neq 0$, by Definition 5, T_9.

Proof of T_{16}. $a\neq 0$, $b\neq 0$, by Hypothesis. Hence, $(a/b)_1=1/(a/b)$ $=(b/b)/(a/b)=b/a$, by Definition 5, Definition 4, T_{11}.

Proof of T_{17}. If $a\neq 0$, then $0a=0/a_1=0$, by Definition 6, T_9. If $a=0$, then $0a=0$, by Definition 6.

Proof of T_{18}. If $b\neq 0$, then $b_1\neq 0$, by T_{15}; hence, if $b\neq 0$, then $ab=a/b_1$ is in K, by Definition 6, D_3. If $b=0$, then ab is in K, by Definition 6.

Proof of T_{19}. If $a\neq 0$, $b\neq 0$, then $ab=a/b_1=a/(1/b)=b/(1/a)=b/a_1=ba$, by Definition 6, Definition 5, D_2, Definition 5, Definition 6. If $a=0$, or $b=0$, then $ab=0=ba$, by T_{17}, Definition 6.

Proof of T_{20}. If $a\neq 0$, $b\neq 0$, $c\neq 0$, then $a(bc)=(bc)a=(b/c_1)/a_1=(b/a_1)/c_1$ $=(ba)c=(ab)c$, by T_{19}, Definition 6, T_{12}, Definition 6, T_{19}. If $a=0$, or $b=0$, or $c=0$, then $a(bc)=0=(ab)c$, by T_{17}, Definition 6.

Proof of T_{21}. $a(b/a)=a/(b/a)_1=a/(a/b)=b$, by Definition 6, T_{16}, T_{10}.

Proof of T_{22}. We first prove the Lemma $ab'=(ab)'$. The Lemma is true because, if $a\neq 0$, then $ab'=b'a=b'/a_1=(0-b)/a_1=0/a_1-b/a_1=0-b/a_1$ $=(b/a_1)'=(ba)'=(ab)'$, by T_{19}, Definition 6, Definition 2, D_6, T_9, Definition 2, Definition 6, T_{19}; if $a=0$, then $ab'=0=0-0=0-0b=(0b)'=(ab)'$, by T_{17}, Definition 1, T_{17}, Definition 2, hypothesis. If now $a\neq 0$, then $a(b+c)=(b+c)a$

$=(b-c')a=(b-c')/a_1=b/a_1-c'/a_1=ba-c'a=ab-ac'=ab-(ac)'=ab+ac$, by T_{19}, Definition 3, Definition 6, D_6, Definition 6, T_{19}, Lemma, Definition 3; if $a=0$, then $a(b+c)=0=0-(0-0)=0-0'=0+0=0b+0c=ab+ac$, by T_{17}, S_1, Definition 2, Definition 3, T_{17}, hypothesis.

6. **Sufficiency of Postulates** (F) **for fields.** From propositions N, T_5–T_8, T_{18}–T_{22} we see that *Postulates* (F) *are sufficient for a field.**

7. **Postulates** (F′), (F″) **for fields.** Following are two other sets of postulates for fields:

(F′) N, T_1, S_2, S_3, D_1–D_6.
(F″) N, T_1, T_3, S_3, D_1–D_6.

In T_1 is understood the condition *if a, b, and the combinations indicated are in K*. In T_3 is understood the condition *if a, b, c, and the combinations indicated are in K*.

It is clear that the sets (F′) and (F″) are each consistent and necessary for a field.

The postulates in each of the sets (F′) and (F″) are independent. Independence-systems for (F′) are the same as those for (F), given by the table of §3, except that systems $\bar{S}_1$, $\bar{S}_2$ are replaced respectively by (α), (β) following:

(α) $K=0,1; a-b=0; a/b=a+(0\div b)$.
(β) $K=0,1; a-b=b; a/b=0\div 0$.

Independence-systems for (F″) are the same as those for (F), except that systems $\bar{S}_1$, $\bar{S}_2$ are replaced respectively by (γ), (δ) following:

(γ) $K=0,1; a-b=a; a/b=0\div 0$.
(δ) $K=0,1; a-b=b; a/b=0\div 0$.

The proof of the sufficiency of (F′), (F″) for fields is left to the reader.

8. **Postulates for abelian groups.** Corresponding to (F), (F′), (F″) we have respectively (A), (A′), (A″) following as postulate sets for (non-trivial) abelian groups:

(A) N, S_1, S_2, S_3.
(A′) N, T_1, S_2, S_3.
(A″) N, T_1, T_3, S_3.

Clearly, (A), (A′), (A″) are each consistent and necessary for abelian groups.

Independence-systems for (A), (A′), (A″) respectively are the independence-systems for the corresponding fields confined to K, $-$.

* Compare E. V. Huntington, *Definitions of a field by sets of independent postulates*, these Transactions, vol. 4 (1903), p. 33.

The sufficiency of (A) for abelian groups follows from N, T_5–T_8.* The proof of the sufficiency of (A′) and of (A″) is left to the reader.†

* Compare E. V. Huntington, *Two definitions of an abelian group by sets of independent postulates*, these Transactions, vol. 4 (1903), p. 27.

† Since the above was written, two papers by David G. Rabinow have appeared: *Independent sets of postulates for abelian groups and fields in terms of the inverse operations*, American Journal of Mathematics, vol. 59 (1937), pp. 211–224; *Note on the definition of fields by independent postulates in terms of the inverse operations*, ibid., pp. 385–392. In the first of these papers the author recapitulates his results in an earlier paper, entitled *Independent set of postulates for a group in terms of the inverse operation*, offered to the Bulletin of the Society about May 6, 1936. This is the paper of Rabinow's to which I referred in my first footnote. The results in this paper differ little from those in Ward's paper (cited above). As far as I know, the paper has not yet been published. Rabinow's two Journal papers are closely related to mine. (In the first paper, the postulates for abelian groups are precisely my set (A″), without my Postulate N and without the restrictions on the elements in my T_1, T_3.) In neither paper is there reference to Ward's paper or to mine (Abstract 42-5-133, Bulletin of the American Mathematical Society).

THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIF.

THE STIELTJES TRANSFORM*

BY

D. V. WIDDER

Introduction. The Stieltjes transform is defined by the equation

$$f(x) = \int_0^\infty \frac{d\alpha(t)}{x+t} = \lim_{R\to\infty} \int_0^R \frac{d\alpha(t)}{x+t}. \tag{1}$$

We assume that $\alpha(t)$ is of bounded variation in $(0, R)$ for every positive R and that the limit (1) exists. If $\alpha(t)$ is the integral of a function $\phi(t)$, we obtain the special case

$$f(x) = \int_0^\infty \frac{\phi(t)}{x+t} dt. \tag{2}$$

In this paper we discuss two distinct but related questions, the inversion problem and the solubility problem. In the former we assume that $f(x)$ is a function which admits the representation (1) or (2), and seek to determine $\alpha(t)$ or $\phi(t)$ from $f(x)$. In the latter we seek necessary and sufficient conditions on $f(x)$ that it should have the representation (1) or (2).

A solution of the inversion problem was given by Stieltjes† himself by means of contour integration. His result was

$$\frac{\alpha(t+) + \alpha(t-)}{2} = \lim_{\gamma\to 0} R \frac{1}{\pi i} \int_{-t-i\gamma}^{-i\gamma} f(s) ds,$$

where the symbol R means "real part of."

It is our purpose to obtain a real inversion formula, one depending only on a knowledge of $f(x)$ and its derivatives on the positive real axis. One may conjecture the existence of such a formula by noting that the Stieltjes transform is the product of two Laplace transforms. That is,

$$f(x) = \int_0^\infty e^{-xu} du \int_0^\infty e^{-ut} d\alpha(t).$$

But a Laplace transform admits of two types of inversion, one by contour integration and one by use of the successive derivatives of $f(x)$ on the positive real axis.‡ As we showed in the paper cited, these two inversions are

* Presented to the Society, December 31, 1936; received by the editors March 1, 1937.

† *Collected works of T. J. Stieltjes*, vol. 2, p. 473.

‡ D. V. Widder, *The inversion of the Laplace integral and the related moment problem*, these Transactions, vol. 36 (1934), p. 107.

analogous to the two classical determinations of the coefficients of a power series, one by Cauchy's integral, the other by Taylor's series.

A real inversion of the integral (2) has been found recently by Paley and Wiener* in case $\phi(t)$ and $[\phi(t)]^2$ are integrable in the interval $(0, \infty)$. The result is

$$\text{(3)} \qquad \underset{m\to\infty}{\text{l.i.m.}}\ \frac{1}{\pi t^{1/2}} \sum_{n=0}^{m} \frac{(-1)^n}{(2n)!}\left(\pi t \frac{d}{dt}\right)^{2n} [t^{1/2} f(t)].$$

The formula obtained in the present paper seems, at first sight, totally unrelated, but we shall show later that this is not the case. It is

$$L_{k,t}[f(x)] = \frac{(-t)^{k-1}}{k!(k-2)!} \frac{d^{2k-1}}{dt^{2k-1}} [t^k f(t)].$$

This is a linear differential operator of order $2k-1$. With no restrictions on $\phi(t)$ beyond those necessary to make (2) converge we show that

$$\lim_{k\to\infty} L_{k,t}[f(x)] = \phi(t)$$

for almost all positive t. Further, we prove that

$$\lim_{k\to\infty} \int_0^t L_{k,u}[f(x)]du = \frac{\alpha(t+) + \alpha(t-)}{2} - \alpha(0+)$$

for all positive t.

The method employed is the same as that used earlier by the author for the Laplace transform. That is, one employs a known method for discussing the asymptotic behavior of an integral of the form

$$\int_0^\infty [g(t)]^k \phi(t) dt$$

as k becomes infinite. In the present case

$$g(t) = \frac{t}{(x+t)^2},$$

a function which has a single maximum at $t=x$. We observe that the fundamental solutions of the linear differential expression $L_{k,t}[f(x)]$ are

$$f(x) = x^n \qquad (n = -k, -k+1, -k+2, \cdots, -1, 0, 1, \cdots, k-2).$$

Hence we may say that the Stieltjes transform is inverted by the linear differ-

* R. E. A. C. Paley and N. Wiener, *Fourier transforms in the complex domain,* American Mathematical Society Colloquium Publications, vol. 19, 1934.

ential operator of infinite order which has all integral powers of x as its fundamental solutions.

We solve the solubility problem by obtaining necessary and sufficient conditions that $f(x)$ should have the representation (1) or (2) with $\alpha(t)$ or $\phi(t)$ belonging to certain familiar classes. The most important classes considered are $\alpha(t)$ of bounded variation or non-decreasing, $\phi(t)$ of class $L^p;(p \geqq 1)$ or bounded. For example, we prove that $f(x)$ has the form (1) with $\alpha(t)$ non-decreasing if and only if

$$f(x) \geqq 0, \qquad (-1)^k[x^k f(x)]^{(2k-1)} \geqq 0 \qquad (x > 0;\ k = 1, 2, \cdots),$$
$$f(\infty) = 0.$$

This is the analogue of Bernstein's theorem on completely monotonic functions. The case of the integral (2) with $\phi(t)$ belonging to L^2 was treated by Paley and Wiener in the Colloquium lectures cited above.

A formula for computing the saltus of $\alpha(t)$ at a point of discontinuity is also obtained. In fact we show that

$$\lim_{k\to\infty} 2t\left(\frac{\pi}{k}\right)^{1/2} L_{k,t}[f(x)] = \alpha(t+) - \alpha(t-) \qquad (t > 0).$$

An inversion formula for the generalized Stieltjes transform

$$f(x) = \int_0^\infty \frac{d\alpha(t)}{(x+t)^\rho} \qquad (\rho > 0)$$

is also obtained.

The final section of the paper shows the relation between the Paley-Wiener operator and $L_{k,t}[f(x)]$. The former can be written symbolically as

$$(\cos \pi\mathcal{D})[t^{1/2}f(t)],$$

where

$$\mathcal{D} = t\frac{d}{dt}\cdot$$

For, the finite series (3) is clearly a section of the infinite power series for $\cos \pi\mathcal{D}$. We show that $L_{k,t}[f(x)]$ is essentially a section of the familiar infinite product expansion of the cosine, so that symbolically the operators are equivalent. It should be observed that the Paley-Wiener operator is applicable only to the case in which $\phi(t)$ belongs to L^2 (at least in so far as proofs have yet been given), whereas the operator of the present paper is not so restricted.

1. **General properties.** Let $\alpha(t)$ be a complex function of the real variable t of bounded variation in the interval $0 \leqq t \leqq R$ for every positive R. Such a function is said to be normalized if

$$\alpha(0) = 0,$$
$$\alpha(t) = \frac{\alpha(t+) + \alpha(t-)}{2} \qquad (t > 0).$$

We assume throughout that $\alpha(t)$ has these properties. It is clear that the integral

$$\int_0^R \frac{d\alpha(t)}{s+t}$$

exists for every complex $s = \sigma + i\tau$ not on the negative real axis, $\tau = 0$, $\sigma \leqq 0$, which we shall henceforth denote by D. Set

$$f(s) = \int_0^\infty \frac{d\alpha(t)}{s+t} = \lim_{R\to\infty} \int_0^R \frac{d\alpha(t)}{s+t} \tag{1.1}$$

whenever the indicated limit exists. Then the improper integral (1.1) is said to converge. We show at once that the region of convergence of (1.1) is the whole complex plane less the ray D.

THEOREM 1.1. *If the integral* (1.1) *converges for a point s_0 not on D, it converges for every such point.*

For, set

$$\beta(0) = 0, \quad \beta(t) = \int_0^t \frac{d\alpha(u)}{s_0 + u} \qquad (t > 0).$$

Then for s not on D

$$\int_0^R \frac{d\alpha(t)}{s+t} = \int_0^R \frac{s_0+t}{s+t}\, d\beta(t) = \beta(R)\frac{s_0+R}{s+R} + (s_0 - s)\int_0^R \frac{\beta(t)}{(s+t)^2}\, dt.$$

Since $\beta(R)$ approaches a limit as R becomes infinite, it is clear that the integral

$$\int_0^\infty \frac{\beta(t)}{(s+t)^2}\, dt$$

converges absolutely and that (1.1) converges. Moreover,

$$\int_0^\infty \frac{d\alpha(t)}{s+t} = \int_0^\infty \frac{d\alpha(t)}{s_0+t} + (s_0 - s)\int_0^\infty \frac{\beta(t)}{(s+t)^2}\, dt. \tag{1.2}$$

We observe that (1.1) may converge without converging absolutely as the example.

$$f(s) = \sum_{n=0}^{\infty} \frac{(-1)^n}{s+n}$$

shows. But (1.2) always enables us to replace the given integral by an absolutely convergent one.

At this point we note a contrast with the theory of the Laplace transform. In the latter case a direct integration by parts replaces a conditionally convergent integral by an absolutely convergent one. That this is not always true for the Stieltjes transform is seen by an example. Take

$$\alpha(0) = 0,$$
$$\alpha(t) = (-1)^n \frac{t+2}{\log(t+2)} \qquad (n < t < n+1; n = 0, 1, 2, \cdots).$$

Then

$$\int_0^\infty \frac{|\alpha(t)|}{(t+2)^2} dt = \int_0^\infty \frac{dt}{(t+2)\log(t+2)}$$

clearly diverges. But

$$\int_0^\infty \frac{d\alpha(t)}{t+2} = \int_0^\infty \frac{\alpha(t)}{(t+2)^2} dt = \sum_{n=0}^\infty (-1)^n \log \frac{\log(n+3)}{\log(n+2)},$$

the series converging. For this example, integration by parts replaces a conditionally convergent integral by another with the same property.

COROLLARY 1.11. *If* (1.1) *converges, it converges uniformly in any bounded closed region not containing a point of D.*

COROLLARY 1.12. *If* (1.1) *converges,* $f(s)$ *is analytic at points not on D.*

COROLLARY 1.13. *If* (1.1) *converges,*

$$f^{(k)}(s) = (-1)^k k! \int_0^\infty \frac{d\alpha(t)}{(s+t)^{k+1}} \qquad (k = 0, 1, \cdots).$$

Another very useful result is contained in

THEOREM 1.2. *If* (1.1) *converges, then*

$$\alpha(t) = o(t) \qquad (t \to \infty). \tag{1.3}$$

Let (1.1) converge for $s = s_0$ not on D, and define $\beta(t)$ as in Theorem 1.1. Then

$$\alpha(R) = \int_0^R d\alpha(t) = \int_0^R (t + s_0)d\beta(t).$$

But

$$\int_0^R td\beta(t) = R\beta(R) - \int_0^R \beta(t)dt,$$

$$\lim_{R\to\infty} \frac{1}{R}\int_0^R td\beta(t) = \beta(\infty) - \beta(\infty) = 0.$$

Also

$$\lim_{R\to\infty} \frac{1}{R}\int_0^R s_0 d\beta(t) = 0,$$

so that (1.3) is established.

COROLLARY 1.21. *If* (1.1) *converges, then*

$$f(s) = \int_0^\infty \frac{\alpha(t)}{(s+t)^2}\, dt.$$

It is important to note that the converse of Theorem 1.2 is false. Thus (1.3) holds if

$$\alpha(t) = \int_0^t \frac{du}{\log(u+2)} \qquad (t \geqq 0).$$

Yet for this definition of $\alpha(t)$ the integral (1.1) diverges.

However, it is easily seen that if

$$\alpha(t) = O(t^{1-\delta}) \qquad (\delta > 0, t \to \infty),$$

then (1.1) converges.

The relation between the Laplace and Stieltjes transforms is made precise in

THEOREM 1.3. *If the integral* (1.1) *converges, then*

$$f(s) = \int_{0+}^\infty e^{-st}\phi(t)dt \qquad (\sigma > 0), \tag{1.4}$$

*where**

$$\phi(t) = \int_0^\infty e^{-tu}d\alpha(u) \qquad (t > 0). \tag{1.5}$$

* The notation employed in (1.4) means that $\int_{0+}^\infty g(t)dt = \lim_{\epsilon\to 0} \int^\infty g(t)dt$.

For, by Theorem 1.2 $\alpha(u)=o(u)$ when u becomes infinite, so that (1.5) converges uniformly* in the closed interval $\epsilon\leqq t\leqq R$ for arbitrary positive numbers ϵ and R. Hence

$$\int_\epsilon^R e^{-st}dt\int_0^\infty e^{-tu}d\alpha(u) = \int_0^\infty \frac{e^{-(s+u)\epsilon} - e^{-(s+u)R}}{s+u}\,d\alpha(u).$$

If s is any point not on the ray D, the integral

$$\int_0^\infty \frac{e^{-\epsilon u}}{s+u}\,d\alpha(u)$$

clearly converges, so that

$$\int_\epsilon^R e^{-st}dt\int_0^\infty e^{-tu}d\alpha(u) = e^{-s\epsilon}\int_0^\infty \frac{e^{-\epsilon u}}{s+u}\,d\alpha(u) - e^{-sR}\int_0^\infty \frac{e^{-uR}}{s+u}\,d\alpha(u). \tag{1.6}$$

The first integral on the right-hand side converges uniformly† (s being fixed) in the interval $0\leqq\epsilon<\infty$, and hence approaches $f(s)$ as ϵ approaches zero. Moreover,†

$$\lim_{R\to\infty}\int_0^\infty \frac{e^{-uR}}{s+u}\,d\alpha(u) = \lim_{t\to 0+}\int_0^t \frac{d\alpha(u)}{s+u} = \frac{\alpha(0+)}{s},$$

so that the last term of (1.6) approaches zero with $1/R$ if $\sigma>0$. Our result is consequently established by allowing ϵ to approach zero and R to become infinite.

Note that the inequality $\sigma>0$ in (1.4) cannot be replaced by $\sigma\geqq 0$ as the example $f(s)=1/s$ shows. We observe also that the converse of Theorem 1.3 is not true. That is, the integrals (1.4) and (1.5) may converge in the range specified without having $f(s)$ represented in the form (1.1). For example, take

$$\phi(t) = \frac{1}{(1+e^{-t})^2} = \sum_{n=0}^\infty (-1)^n(n+1)e^{-nt} \qquad (t>0),$$

$$f(s) = \int_0^\infty \frac{e^{-st}}{(1+e^{-t})^2}\,dt \qquad (\sigma>0).$$

The integral (1.1) becomes

$$\sum_{n=0}^\infty (-1)^n\frac{(n+1)}{s+n},$$

---

* See footnote on p. 12.

† For the results regarding the Laplace transform which are here employed see, for example, D. V. Widder, *A generalization of Dirichlet's series and of Laplace's integrals by means of a Stieltjes integral*, these Transactions, vol. 31 (1929), p. 694.

a series which diverges for all s. In this connection we may prove

THEOREM 1.4. *If $\alpha(u)$ is such that the integral* (1.5) *converges for $t \geqq 0$, then the function $f(s)$ defined by* (1.4) *also has the representation* (1.1) *for all s not on D.*

For, in this case, $\alpha(u)$ is necessarily bounded and we may apply Theorem 1.3.

A similar result is contained in

THEOREM 1.5. *If $\alpha(u)$ is non-decreasing and such that* (1.4) *and* (1.5) *converge for $\sigma > 0$, $t > 0$ respectively, then the function $f(s)$ defined by* (1.4) *has the representation* (1.1) *for all s not on D.*

The proof is easily supplied. We turn next to the uniqueness theorem.

THEOREM 1.6. *If the normalized function $\alpha(t)$ is such that*

$$\int_0^\infty \frac{d\alpha(t)}{s_0 + t + nl} = 0 \qquad (l > 0, n = 0, 1, 2, \cdots),$$

then

$$\alpha(t) = 0 \qquad (0 \leqq t < \infty).$$

This follows from Theorem 1.3 and from Lerch's uniqueness theorem for Laplace integrals.*

We conclude this section with a proof of

THEOREM 1.7. *If $\alpha(t)$ is a real non-decreasing function for which the point $t = 0$ is a point of increase and for which the integral*

$$f(s) = \int_0^\infty \frac{d\alpha(t)}{s + t}$$

converges, then $f(s)$ has a singularity at $s = 0$.

For, suppose the contrary. Then the series

$$f(s) = \sum_{n=0}^\infty f^{(n)}(1) \frac{(s-1)^n}{n!}$$

converges for some point on the negative real axis, say $s = -\epsilon$, and

$$f(-\epsilon) = \sum_{n=0}^\infty (-1)^n f^{(n)}(1) \frac{(\epsilon + 1)^n}{n!}.$$

Applying Corollary 1.13 we obtain

* The usual proof of this theorem can easily be extended to include Cauchy-values of Laplace integrals: $\int_{0+}^\infty e^{-st} d\alpha(t)$.

$$f(-\epsilon) = \sum_{n=0}^{\infty} \int_0^{\infty} \frac{(\epsilon+1)^n}{(t+1)^{n+1}} d\alpha(t)$$
$$= \sum_{n=0}^{\infty} (n+1) \int_0^{\infty} \frac{(\epsilon+1)^n}{(t+1)^{n+2}} \alpha(t)dt.$$

This series dominates the series

$$\sum_{n=0}^{\infty} (n+1) \int_{\epsilon}^{\infty} \frac{(\epsilon+1)^n}{(t+1)^{n+2}} \alpha(t)dt, \tag{1.7}$$

so that the latter also converges. Since the integrand is non-negative, and since the series

$$\sum_{n=0}^{\infty} (n+1) \frac{(\epsilon+1)^n}{(t+1)^{n+2}} = \frac{1}{(t-\epsilon)^2}$$

converges for $t>\epsilon$, we may interchange integral and summation symbols in (1.7). That is, the integral

$$\int_{\epsilon}^{\infty} \frac{\alpha(t)}{(t-\epsilon)^2} dt \tag{1.8}$$

must converge. But since $t=0$ is a point of increase for $\alpha(t)$, it follows that $\alpha(\epsilon+)>0$ and

$$\lim_{t\to\epsilon+} \frac{\alpha(t)}{t-\epsilon} = +\infty.$$

Hence (1.8) can not converge. The assumption that $f(s)$ is analytic at $s=0$ is untenable. That is, $s=0$ is a singularity of $f(s)$.

2. **Inversion in a special case.** The results of the previous section enable us to restrict attention to the real variable $s=x$. In fact we shall even assume that $\alpha(t)$ is a real function. The loss of generality thus involved is trivial. The reader who has need of results for complex functions $\alpha(t)$ has only to apply the theorems proved to the real and imaginary parts of $\alpha(t)$ separately.

We introduce a functional operator by

DEFINITION 2.1. *An operator $L_{k,t}[f(x)]$ is defined by the equations*

$$L_{0,t}[f(x)] = c_0 f(t),$$
$$L_{k,t}[f(x)] = c_k(-t)^{k-1}[t^k f(t)]^{(2k-1)} \quad (k = 1, 2, 3, \cdots),$$

where

$$c_0 = c_1 = 1, \quad c_k = \frac{1}{k!(k-2)!} \quad (k = 2, 3, 4, \cdots).$$

Obviously the operator can be applied only to functions which possess derivatives of order $(2k-1)$. It becomes of interest only when k is allowed to become infinite, so that we shall be applying it only to functions which possess derivatives of all orders. Our first application will be to the function (1.1) where $\alpha(t)$ is a step-function defined as follows:

$$\alpha(t) = \begin{cases} 0 & (0 \leqq t < a), \\ 1 & (a < t < \infty), \end{cases}$$

$$\alpha(a) = \tfrac{1}{2}.$$

In this case

$$f(x) = \int_0^\infty \frac{d\alpha(t)}{x+t} = \frac{1}{x+a}\,.$$

Simple computation gives

$$L_{k,t}[f(x)] = d_k \frac{t^{k-1}a^k}{(t+a)^{2k}} \qquad (k = 1, 2, 3, \cdots),$$

$$d_k = (2k-1)!c_k.$$

We now prove

THEOREM 2.1. *If $a>0$, $t>0$, then*

$$\lim_{k\to\infty} \int_0^t L_{k,u}\left[\frac{1}{x+a}\right] du = \begin{cases} 0 & (0 \leqq t < a), \\ \tfrac{1}{2} & (t = a), \\ 1 & (1 < t < \infty). \end{cases}$$

That is, the operator $L_{k,t}[f(x)]$ serves to invert the integral (1.1) at least in this special case. Set

$$H_k(t) = d_k \int_0^t \frac{u^{k-1}}{(u+1)^{2k}}\, du.$$

Then

$$\int_0^t L_{k,u}\left[\frac{1}{x+a}\right] du = H_k\left(\frac{t}{a}\right),$$

so that we need prove only

$$\lim_{k\to\infty} H_k(t) = \begin{cases} 0 & (0 \leqq t < 1), \\ \tfrac{1}{2} & (t = 1), \\ 1 & (1 < t < \infty). \end{cases}$$

If $0 \leqq t < 1$ we have

$$0 \leqq H_k(t) < d_k \left[\frac{t}{(t+1)^2}\right]^{k-1}. \tag{2.1}$$

Since

$$\lim_{k\to\infty} \frac{(2k-1)(2k-2)}{k(k-2)} \frac{t}{(t+1)^2} = \frac{4t}{(t+1)^2} < 1,$$

it follows that the extreme right-hand member of (2.1), and hence also $H_k(t)$, tends to zero with $1/k$.

If $1 < t < \infty$, we have by use of the B-function

$$\frac{k-1}{k} - H_k(t) = d_k \int_t^\infty \frac{u^{k-1}}{(u+1)^{2k}} du.$$

But

$$0 < d_k \int_t^\infty \frac{u^{k-1}}{(u+1)^{2k}} du < d_k \left[\frac{t}{(t+2)^2}\right]^{k-1},$$

so that in this case

$$\lim_{k\to\infty} H_k(t) = 1.$$

Finally, if $t = 1$,

$$H_k(1) = \frac{k-1}{k} - d_k \int_1^\infty \frac{u^{k-1}}{(u+1)^{2k}} du = \frac{k-1}{k} - d_k \int_0^1 \frac{u^{k-1}}{(u+1)^{2k}} du$$
$$= \frac{k-1}{k} - H_k(1)$$

by an obvious change of variable. Hence

$$H_k(1) = \frac{1}{2} \frac{k-1}{k} \to \frac{1}{2} \qquad (k \to \infty).$$

This completes the proof of the theorem.

3. **The inversion of the general Stieltjes integral.** Before proceeding to the general case we need to prove the following simple, but extremely useful lemma:

LEMMA 3.11. *If $f(x)$ has a derivative of order $(2k-1)$, then*

$$x^{k-1}\left[x^k f(x)\right]^{(2k-1)} = \left[x^{2k-1} f^{(k-1)}(x)\right]^{(k)}.$$

The proof consists merely in computing both sides of the equation by Leibniz's rule. In each case we obtain

$$\sum_{p=0}^{k} \frac{(2k-1)!k!}{(2k-p-1)!p!(k-p)!} f^{(2k-p-1)}(x) x^{2k-p-1}.$$

We shall also need

LEMMA 3.12. *If* (1.1) *converges, then*

$$\lim_{x\to 0+} d_k x^k \int_0^\infty \frac{t^{k-1}\alpha(t)}{(x+t)^{2k}}\,dt = \alpha(0+)\,\frac{k-1}{k} \qquad (k = 2, 3, \cdots).$$

For, if ϵ is a given positive number, we determine $\delta(\epsilon)$ such that

$$|\alpha(t) - \alpha(0+)| < \epsilon \qquad 0 \leqq t \leqq \delta(\epsilon).$$

Then

$$\left| d_k x^k \int_0^\infty \frac{t^{k-1}[\alpha(t) - \alpha(0+)]}{(x+t)^{2k}}\,dt \right| \leqq \epsilon d_k x^k \int_0^{\delta(\epsilon)} \frac{t^{k-1}}{(x+t)^{2k}}\,dt \tag{3.1}$$
$$+ \left| d_k x^k \int_{\delta(\epsilon)}^\infty \frac{t^{k-1}}{(x+t)^{2k}} [\alpha(t) - \alpha(0+)]\,dt \right|,$$

$$\limsup_{x\to 0+} \left| d_k x^k \int_0^\infty \frac{t^{k-1}[\alpha(t) - \alpha(0+)]}{(x+t)^{2k}}\,dt \right| \leqq \epsilon\,\frac{k-1}{k} < \epsilon,$$

the second term on the right-hand side of (3.1) clearly approaching zero with x. Hence

$$\lim_{x\to 0+} d_k x^k \int_0^\infty \frac{t^{k-1}}{(x+t)^{2k}}\,\alpha(t)\,dt = d_k \alpha(0+) x^k \int_0^\infty \frac{t^{k-1}}{(x+t)^{2k}}\,dt = \alpha(0+)\,\frac{k-1}{k}.$$

In a similar way we can prove

LEMMA 3.13. *If* $\alpha(\infty)$ *exists, then*

$$\lim_{x\to \infty} d_k x^k \int_0^\infty \frac{t^{k-1}}{(x+t)^{2k}}\,\alpha(t)\,dt = \alpha(\infty)\,\frac{k-1}{k} \qquad (k = 2, 3, 4, \cdots).$$

By use of these results we can now establish

THEOREM 3.1. *If the integral* (1.1) *converges, then*

$$\lim_{k\to\infty} \int_0^t L_{k,u}[f(x)]\,du = \alpha(t) - \alpha(0+) \qquad (t > 0).$$

We begin by writing the integral (1.1) in the form

$$f(x) = \int_0^\infty \frac{\alpha(t)}{(x+t)^2}\,dt,$$

which we are enabled to do by Corollary 1.21. By Lemma 3.11 we have

$$\int L_{k,u}[f(x)]\,du = c_k[u^{2k-1} f^{(k-1)}(u)]^{(k-1)}.$$

But simple computation gives

$$(-1)^{k-1}c_k\left[u^{2k-1}f^{(k-1)}(u)\right]^{(k-1)} = d_k u^k \int_0^\infty \frac{y^{k-1}\alpha(y)}{(u+y)^{2k}}\,dy.$$

Hence by Lemma 3.12

$$\int_0^t L_{k,u}[f(x)]du = d_k t^k \int_0^\infty \frac{y^{k-1}\alpha(y)}{(t+y)^{2k}}\,dy - \frac{k-1}{k}\,\alpha(0+).$$

Consequently, it remains only to show that

$$\lim_{k\to\infty} d_k t^k \int_0^\infty \frac{y^{k-1}\alpha(y)}{(t+y)^{2k}}\,dy = \alpha(t) \qquad (t>0). \tag{3.2}$$

Set $y/t=v$. The integral in question becomes

$$d_k \int_0^\infty \frac{v^{k-1}\alpha(tv)}{(1+v)^{2k}}\,dv.$$

Further, set

$$\psi_t(v) = \begin{cases} \alpha(t-) & (0<v<1), \\ \alpha(t) & (v=1), \\ \alpha(t+) & (1<v<\infty). \end{cases}$$

Then*

$$d_k \int_0^\infty \frac{v^{k-1}}{(1+v)^{2k}}\,\psi_t(v)dv = [\alpha(t+)+\alpha(t-)]H_k(1).$$

Hence by Theorem 2.1

$$\lim_{k\to\infty} d_k \int_0^\infty \frac{v^{k-1}}{(1+v)^{2k}}\,\psi_t(v)dv = \frac{\alpha(t+)+\alpha(t-)}{2} = \alpha(t).$$

Set

$$\beta_t(v) = \alpha(tv) - \psi_t(v).$$

This function is continuous at $t=1$ and vanishes there. It remains only to show that

$$\lim_{k\to\infty} d_k \int_0^\infty \frac{v^{k-1}}{(1+v)^{2k}}\,\beta_t(v)dv = 0.$$

Given $\epsilon>0$, we determine $\delta(\epsilon)$ such that

* For the definition of $H_k(t)$ see §2.

$$|\beta_t(v)| < \epsilon \qquad |v-1| \leqq \delta(\epsilon).$$

Then

$$\left| d_k \int_0^\infty \frac{v^{k-1}}{(1+v)^{2k}} \beta_k(v)dv \right| \leqq O(1)H_k(1-\delta) + \epsilon\{H_k(1+\delta) - H_k(1-\delta)\} \tag{3.3}$$
$$+ O(1)\left\{\frac{k-2}{k-1} - H_{k-1}(1+\delta)\right\}.$$

In obtaining the last term on the right-hand side of (3.3), we have used the obvious fact that

$$\beta_t(v) = O(v) \qquad (v \to \infty).$$

Letting k become infinite in (3.3) and making use of Theorem 2.1, we have

$$\limsup_{k\to\infty} \left| d_k \int_0^\infty \frac{v^{k-1}}{(1+v)^{2k}} \beta_t(v)dv \right| \leqq \epsilon,$$

from which our result follows at once.

COROLLARY 3.11. *If* (1.1) *converges and if* $\alpha(t)$ *is continuous in* (a, b), *then*

$$\lim_{k\to\infty} \int_0^t L_{k,u}[f(x)]du = \alpha(t) - \alpha(0+)$$

uniformly in the interval $\alpha \leqq t \leqq \beta$, *where*

$$a < \alpha < \beta < b \qquad (a > 0),$$
$$a \leqq \alpha < \beta < b \qquad (a = 0).$$

To prove this one has only to show that

$$\lim_{k\to\infty} d_k t^k \int_0^\infty \frac{y^{k-1}}{(t+y)^{2k}} [\alpha(y) - \alpha(t)]dy$$
$$= \lim_{k\to\infty} d_k \int_0^\infty \frac{y^{k-1}}{(t+y)^{2k}} [\alpha(ty) - \alpha(t)]dy = 0$$

uniformly in (α, β). If we note that

$$\alpha(ty) - \alpha(t) = o(1) \qquad (y \to 1)$$

uniformly in t for t in (α, β), the proof of this proceeds as for Theorem 3.1.

COROLLARY 3.12. *If* $\alpha(\infty)$ *exists, then*

$$\lim_{k\to\infty} \int_0^\infty L_{k,u}[f(x)]du = \alpha(\infty) - \alpha(0+).$$

COROLLARY 3.13. *If*

$$\alpha(t) = O(t^\rho) \qquad (t \to \infty)$$

for some positive value of ρ, *then equation* (3.2) *holds.*

4. **The Lebesgue integral.** In this section we consider the special case of (1.1) in which

$$\alpha(t) = \int_0^t \phi(u)du,$$

the function $\phi(u)$ being integrable in $(0, R)$ for every positive R and of such a nature that (1.1) converges. We then have

$$f(x) = \int_0^\infty \frac{\phi(t)}{x+t} dt. \tag{4.1}$$

We now show that the operator $L_{k,t}[f(x)]$ serves to invert this integral also.

THEOREM 4.1. *If t is a point of the Lebesgue set for the function $\phi(u)$, and if* (4.1) *converges, then*

$$\lim_{k\to\infty} L_{k,t}[f(x)] = \phi(t). \tag{4.2}$$

By the Lebesgue set we mean those points t for which

$$\int_0^h |\phi(u+t) - \phi(t)| du = o(h) \qquad (h \to 0).$$

Direct computation gives

$$L_{k,t}[f(x)] = d_k t^{k-1} \int_0^\infty \frac{u^k}{(t+u)^{2k}} \phi(u)du.$$

We have to show that

$$\lim_{k\to\infty} d_k t^{k-1} \int_0^\infty \frac{u^k}{(t+u)^{2k}} [\phi(u) - \phi(t)]du$$
$$= \lim_{k\to\infty} d_k \int_0^\infty \frac{u^k}{(u+1)^{2k}} [\phi(tu) - \phi(t)]du = 0.$$

Set

$$\beta(y) = \int_1^y [\phi(ut) - \phi(t)]du.$$

Since t is a point of the Lebesgue set, we have

$$\beta(y) = o(|1-y|) \qquad (y \to 1). \tag{4.3}$$

Then the integral in question becomes

$$I_k = d_k \int_0^\infty \frac{u^k}{(u+1)^{2k}}\, d\beta(u) = d_k k \int_0^\infty \frac{(u-1)u^{k-1}}{(u+1)^{2k+1}} \beta(u)du.$$

If we note that

$$kd_k \int_{1-\delta}^{1+\delta} \frac{(u-1)^2 u^{k-1}}{(u+1)^{2k+1}}\, du < kd_k \int_0^\infty \frac{(u-1)^2 u^{k-1}}{(u+1)^{2k+1}}\, du = 1$$

and take account of (4.3), we have, by a method similar to that used in obtaining (3.2),

$$| I_k | \leqq O(1)H_k[1 - \delta(\epsilon)] + \epsilon + O(1)\left\{\frac{k-2}{k-1} - H_{k-1}[1 + \delta(\epsilon)]\right\},$$

$$\limsup_{k\to\infty} | I_k | \leqq \epsilon,$$

$$\lim_{k\to\infty} I_k = 0.$$

This proves the theorem.

Corollary 4.11. *Equation* (4.2) *holds for points t at which $\phi(t)$ is continuous.*

Corollary 4.12. *Equation* (4.2) *holds almost everywhere.*

Corollary 4.13. *If* (4.1) *converges and if $\phi(t)$ is continuous in* (a, b), *equation* (4.2) *holds uniformly in $\alpha \leqq t \leqq \beta$ where*

$$a < \alpha < \beta < b.$$

Corollary 4.14. *If $\phi(t+)$ and $\phi(t-)$ exist, then*

$$\lim_{k\to\infty} L_{k,t}[f(x)] = \frac{\phi(t+) + \phi(t-)}{2}.$$

For, set

$$\beta(t, u) = \int_1^t [\phi(yu) - \phi(u)]dy,$$

and note that

$$\beta(t, u) = o(| 1 - t |) \qquad (t \to 1)$$

uniformly in $\alpha \leqq u \leqq \beta$. The proof is now completed by obvious modification of the proof of Theorem 4.1.

At this point we illustrate Theorem 4.1 by an example. Take

$$\phi(t) = \frac{t^{-\delta}}{\Gamma(1-\delta)\Gamma(\delta)} \qquad (0 < \delta < 1).$$

Then

$$f(x) = x^{-\delta}.$$

Simple computation gives

$$L_{k,t}[x^{-\delta}] = \frac{\Gamma(k-\delta+1)\Gamma(k+\delta-1)}{\Gamma(k+1)\Gamma(k-1)} \cdot \frac{t^{-\delta}}{\Gamma(1-\delta)\Gamma(\delta)}.$$

But

$$\frac{\Gamma(k+\alpha)}{\Gamma(k)} \sim k^{\alpha} \qquad (\alpha > 0,\ k \to \infty),$$

so that we can prove directly that

$$\lim_{k\to\infty} L_{k,t}[x^{-\delta}] = \frac{t^{-\delta}}{\Gamma(1-\delta)\Gamma(\delta)} = \phi(t)$$

for all positive t.

5. **The saltus operator.** We now introduce a new operator by the

DEFINITION 5.1. *The operator $l_{k,t}[f(x)]$ is defined by the equation*

$$l_{k,t}[f(x)] = 2t\left(\frac{\pi}{k}\right)^{1/2} L_{k,t}[f(x)].$$

We first apply the operator to the special function

$$f(x) = \int_0^{\infty} \frac{d\psi_t(v)}{x+v} = \frac{\alpha(t+) - \alpha(t-)}{x+1},$$

where $\psi_t(v)$ was defined in §3. Direct computation gives

$$l_{k,u}[f(x)] = 2\left(\frac{\pi}{k}\right)^{1/2} d_k \frac{u^k}{(u+1)^{2k}} [\alpha(t+) - \alpha(t-)].$$

Then by use of Stirling's formula, or otherwise, we prove

LEMMA 5.11.

$$\lim_{k\to\infty} l_{k,u}\left[\frac{1}{x+1}\right] = \begin{cases} 0, & u \neq 1, \\ 1, & u = 1. \end{cases}$$

Hence, the limit of $l_{k,t}[f(x)]$ is the saltus of $\psi_t(u)$ at every point u. This result is general, as we now prove.

THEOREM 5.1. *If* (1.1) *converges, then*

$$\lim_{k\to\infty} l_{k,t}[f(x)] = \alpha(t+) - \alpha(t-) \qquad (t > 0).$$

If we define $\beta(v) = \beta_t(v)$ as in §3 and note that

$$l_{k,t}[f(x)] = 2t^k\left(\frac{\pi}{k}\right)^{1/2} d_k \int_0^\infty \frac{u^k}{(u+t)^{2k}}\, d\alpha(u)$$

$$= 2\left(\frac{\pi}{k}\right)^{1/2} d_k \int_0^\infty \frac{v^k}{(v+1)^{2k}}\, d\alpha(vt),$$

the special example treated above shows us that we need only prove

$$\lim_{k\to\infty} 2\left(\frac{\pi}{k}\right)^{1/2} d_k \int_0^\infty \frac{v^k}{(v+1)^{2k}}\, d\beta(v) = 0,$$

the function $\beta(v)$ being continuous and equal to zero at $v=1$. If we integrate by parts, the integral in question becomes

$$I_k = -2\left(\frac{\pi}{k}\right)^{1/2} d_k \int_0^\infty \beta(v) \frac{d}{dv} \frac{v^k}{(v+1)^{2k}}\, dv.$$

Now note that

$$2\left(\frac{\pi}{k}\right)^{1/2} d_k \int_{1-\delta}^{1+\delta} \left|\frac{d}{dv} \frac{v^k}{(v+1)^{2k}}\right| dv < 2\left(\frac{\pi}{k}\right)^{1/2} d_k \int_0^1 \frac{d}{dv} \frac{v^k}{(v+1)^{2k}}\, dv$$

$$- 2\left(\frac{\pi}{k}\right)^{1/2} d_k \int_1^\infty \frac{d}{dv} \frac{v^k}{(v+1)^{2k}}\, dv = 4\left(\frac{\pi}{k}\right)^{1/2} \frac{d_k}{2^{2k}} = 2l_{k,1}\left[\frac{1}{x+1}\right].$$

Hence, proceeding as in §3, we obtain

$$|I_k| \leqq O(1) l_{k,1-\delta}\left[\frac{1}{x+1}\right] + 2\epsilon l_{k,1}\left[\frac{1}{x+1}\right]$$

$$+ O(1) 2\left(\frac{\pi}{k}\right)^{1/2} d_k \int_{1+\delta}^\infty u \left|\frac{d}{du} \frac{u^k}{(u+1)^{2k}}\right| du.$$

Consequently

$$\limsup_{k\to\infty} |I_k| \leqq 2\epsilon,$$

$$\lim_{k\to\infty} I_k = 0,$$

since

$$2\left(\frac{\pi}{k}\right)^{1/2} d_k \int_{1+\delta}^\infty u \left|\frac{d}{du} \frac{u^k}{(u+1)^{2k}}\right| du$$

$$= (1+\delta)l_{k,1+\delta}\left[\frac{1}{x+1}\right] + 2\left(\frac{\pi}{k}\right)^{1/2} d_k \int_{1+\delta}^{\infty} \frac{u^k}{(u+1)^{2k}}\,du = o(1) \quad (k \to \infty).$$

This completes the proof of the theorem.

The same type of argument enables us to prove the following related result.

THEOREM 5.2. *If* (1.1) *converges and if* $\alpha(t)$ *has right-hand and left-hand derivatives* $\alpha_+'(t)$ *and* $\alpha_-'(t)$ *respectively at a point* t, *then*

$$\lim_{k\to\infty} L_{k,t}[f(x)] = \frac{\alpha_+'(t) + \alpha_-'(t)}{2}.$$

For, if we define

$$\omega(v) = \begin{cases} (v-1)\alpha_-'(t) & 0 \leqq v \leqq 1, \\ (v-1)\alpha_+'(t) & 1 \leqq v < \infty, \end{cases}$$

$$\gamma(v) = \alpha(vt) - \omega(v),$$

it is clear that

$$\gamma(v) = o(|1-v|) \qquad (v \to 1), \tag{5.1}$$

and that*

$$\lim_{k\to\infty} d_k \int_0^{\infty} \frac{v^k}{(v+1)^{2k}}\,d\omega(v) = \frac{\alpha_+'(t) + \alpha_-'(t)}{2},$$

so that we have now to show that the integral

$$I_k = d_k \int_0^{\infty} \gamma(v) \frac{d}{dv} \frac{v^k}{(v+1)^{2k}}\,dv$$

approaches zero with $1/k$. This may be done by use of (5.1) if we note that

$$\begin{aligned} d_k \int_{1-\delta}^{1+\delta} |1-v| \left|\frac{d}{dv}\frac{v^k}{(v+1)^{2k}}\right| dv &\leqq d_k \int_0^1 (1-v)\frac{d}{dv}\frac{v^k}{(v+1)^{2k}}\,dv \\ &\quad + d_k \int_1^{\infty} (1-v)\frac{d}{dv}\frac{v^k}{(v+1)^{2k}}\,dv \\ &= d_k \int_0^{\infty} (1-v)\frac{d}{dv}\frac{v^k}{(v+1)^{2k}}\,dv = d_k \int_0^{\infty} \frac{v^k}{(v+1)^{2k}}\,dv = 1. \end{aligned}$$

COROLLARY 5.21. *If* $\alpha(t)$ *is constant in* (a, b), *then*

---

* This may be conveniently proved by breaking the interval into two parts corresponding to the intervals (0, 1) and $(1, \infty)$ and by using Corollary 4.14.

$$\lim_{k\to\infty} L_{k,t}[f(x)] = 0$$

uniformly for t in (α, β),

$$a < \alpha < \beta < b \qquad (a > 0),$$
$$a \leqq \alpha < \beta < b \qquad (a = 0).$$

Theorem 5.2 becomes of particular interest if $\alpha(t)$ is a function which is not an integral but which has a derivative almost everywhere. The integral (1.1) can not then be put in the form (4.1). Yet the inversion formula has the same effect as if $f(x)$ had the form (4.1) with $\alpha'(t) = \phi(t)$.

6. **Generalizations.** We turn now to a group of theorems which may be regarded as generalizations of our inversion formulas.

THEOREM 6.1. *If* $\alpha(t)$ *is of bounded variation in the interval* $(0, \infty)$, *and if* $\psi(t)$ *is an arbitrary function continuous* in the interval* $0 \leqq t \leqq \infty$, *then*

$$\lim_{k\to\infty} \int_0^\infty \psi(u) L_{k,u}[f(x)]du = \int_0^\infty \psi(u)d\alpha(u) - \alpha(0+)\psi(0). \tag{6.1}$$

Since we have already established that

$$\lim_{k\to\infty} \int_0^t L_{k,u}[f(x)]du = \alpha(t) - \alpha(0+),$$

we have only to take the limit under the integral sign on the left-hand side of (6.1) to obtain our result. This will be permissible by the Helly-Bray theorem† if the functions

$$\int_0^t L_{k,u}[f(x)]du$$

are of uniformly bounded variation in $(0, \infty)$ for $k = 1, 2, 3, \cdots$. This is so under our hypotheses, since

$$\int_0^\infty |L_{k,u}[f(x)]| du = d_k \int_0^\infty u^{k-1}du \int_0^\infty \frac{t^k}{(t+u)^{2k}} |d\alpha(t)| \leqq \int_0^\infty |d\alpha(t)|.$$

THEOREM 6.2. *If* (1.1) *converges, and if* $\psi(t)$ *is continuous in* $(0, R)$, *then*

$$\lim_{k\to\infty} \int_0^R \psi(t) L_{k,t}[f(x)]dt = \int_0^R \psi(t)d\alpha(t) - \psi(0)\alpha(0+).$$

* By this we mean that $\phi(t)$ is continuous for every non-negative value of t and that $\phi(t)$ approaches a limit as t becomes infinite.

† See, for example, G. C. Evans, *The Logarithmic Potential, Discontinuous Dirichlet and Neumann Problems*, American Mathematical Society Colloquium Publications, vol. 6, 1927, p. 15. The infinite intervals in question may be transformed into finite intervals by the transformation $v = e^{-u}$.

Here, we are no longer assuming that $\alpha(t)$ is of bounded variation in $(0, \infty)$. Let δ be an arbitrary positive constant, and set

$$f(x) = \int_0^{R+\delta} \frac{d\alpha(t)}{x+t} + \int_{R+\delta}^{\infty} \frac{d\alpha(t)}{x+t} = f_1(x) + f_2(x).$$

Clearly, Theorem 6.1 is applicable to $f_1(x)$. Let $\psi_R(t)$ be continuous in the infinite interval $(0 \leqq t \leqq \infty)$, coinciding with $\psi(t)$ in $(0, R)$ and constant in (R, ∞). Then

$$\begin{aligned}\lim_{k\to\infty} \int_0^{\infty} \psi_R(t) L_{k,t}[f_1(x)]dt &= \lim_{k\to\infty} \int_0^R \psi(t) L_{k,t}[f_1(x)]dt + \lim_{k\to\infty} \psi(R) \int_R^{\infty} L_{k,t}[f_1(x)]dt \\ &= \int_0^R \psi(t) d\alpha(t) + \psi(R) \int_R^{R+\delta} d\alpha(t) - \psi(0)\alpha(0+).\end{aligned}$$

Making use of Corollary 3.12, we have

$$\begin{aligned}\lim_{k\to\infty} \int_R^{\infty} \psi_R(t) L_{k,t}[f_1(x)]dt &= \lim_{k\to\infty} \psi(R) \int_R^{\infty} L_{k,t}[f_1(x)]dt \\ &= \psi(R)[\alpha(t+R) - \alpha(R)].\end{aligned}$$

On the other hand

$$\lim_{k\to\infty} \int_0^R \psi(t) L_{k,t}[f_2(x)]dt = 0,$$

as one sees by Corollary 5.21. If we combine these results, our theorem is proved.

THEOREM 6.3. *If $\alpha(t)$ has variation $V(R)$ in the interval $0 \leqq t \leqq R$, and if* (1.1) *converges, then*

$$\lim_{k\to\infty} \int_0^R | L_{k,t}[f(x)] | \, dt = V(R) - V(0+). \tag{6.2}$$

It is sufficient to prove the theorem when $\alpha(0+) = \alpha(0) = V(0+) = 0$.

If we set

$$g(x) = \int_0^{\infty} \frac{dV(t)}{x+t},$$

it is clear that

$$| L_{k,t}[f(x)] | \leqq L_{k,t}[g(x)].$$

Hence Theorem 3.1 gives

$$\limsup_{k\to\infty} \int_0^R |L_{k,t}[f(x)]| \, dt \leqq V(R). \tag{6.3}$$

On the other hand, from Theorem 6.2 we have for any function $\psi(t)$ continuous in $(0, R)$

$$\left| \int_0^R \psi(t) d\alpha(t) \right| \leqq \max_{0 \leqq t \leqq R} |\psi(t)| \liminf_{k\to\infty} \int_0^R |L_{k,t}[f(x)]| \, dt.$$

Hence the norm of the linear functional

$$\int_0^R \psi(t) d\alpha(t)$$

which is known to be $V(R)$, is at most

$$\liminf_{k\to\infty} \int_0^R |L_{k,t}[f(x)]| \, dt.$$

That is,

$$V(R) \leqq \liminf_{k\to\infty} \int_0^R |L_{k,t}[f(x)]| \, dt. \tag{6.4}$$

Inequalities (6.3) and (6.4) can not both hold unless (6.2) is true.

COROLLARY 6.31. *If* $V(\infty) < \infty$ *then*

$$\lim_{k\to\infty} \int_0^\infty |L_{k,t}| \, dt = V(\infty) - V(0+).$$

7. **Differentiation and integration of the inversion operator.** Without any restrictions on the integral (1.1) except that it should converge we are able to obtain inversion formulas for the successive integrals of $\alpha(t)$. The following result is seen to be a generalization of Theorem 3.1.

THEOREM 7.1. *If* (1.1) *converges, and if m is any positive integer, then*

$$\lim_{k\to\infty} \int_0^t \frac{(t-u)^m}{m!} L_{k,u}[f(x)] du = \int_0^t du_m \int_0^{u_m} du_{m-1} \int_0^{u_{m-1}} \cdots \int_0^{u_1} d\alpha(u) - \alpha(0+) \frac{t^m}{m!} \cdot$$

The result follows at once from Theorem 6.1 with

$$\psi(u) = \frac{(t-u)^m}{m!} \cdot$$

The successive derivatives of $\alpha(u)$ or $\phi(u)$, when they exist, may be obtained in several ways. We first prove

THEOREM 7.2. *If $\phi(t)$ is of class $C^{(n)}$ in the interval $0 \leq t < \infty$, if*

$$\int_0^\infty \frac{\phi^{(n)}(t)}{x+t}\,dt \tag{7.1}$$

converges, and if $f(x)$ is defined by (4.1), *then*

$$\lim_{k\to\infty} L_{k,t}[(-1)^n f^{(n)}(x)] = \phi^{(n)}(t) \qquad (t > 0).$$

Since (7.1) converges, we have

$$\int_0^t \phi^{(n)}(u)du = o(t) \qquad (t \to \infty).$$

Hence

$$\phi^{(j)}(t) = o(t^{n-j}) \qquad (t \to \infty;\ j = 0, 1, 2, \cdots, n-1), \tag{7.2}$$

so that integration by parts gives us

$$\int_0^\infty \frac{\phi^{(n)}(t)}{x+t}\,dt = -\frac{\phi^{(n-1)}(0)}{x} - \frac{\phi^{(n-2)}(0)}{x^2} - \frac{2\phi^{(n-3)}(0)}{x^3} - \cdots$$

$$- (n-1)!\frac{\phi(0)}{x^n} + n!\int_0^\infty \frac{\phi(t)}{(x+t)^{n+1}}\,dt$$

$$= (-1)^n f^{(n)}(x) - \sum_{j=1}^{n} (j-1)!\frac{\phi^{(n-j)}(0)}{x^j}.$$

The result now follows by use of Corollary 4.11.

COROLLARY 7.21. *If $\alpha(t)$ is of class C^n in $(0, \infty)$, if the integral*

$$\int_0^\infty \frac{\alpha^{(n)}(t)}{x+t}\,dt$$

converges, and if $f(x)$ is defined by (1.1), *then*

$$\lim_{k\to\infty} L_{k,t}[(-1)^{n-1} f^{(n-1)}(x)] = \alpha^{(n)}(t) \qquad (t > 0).$$

A more natural procedure is given in

THEOREM 7.3. *Under the conditions of Theorem* 7.2

$$\lim_{k\to\infty} \frac{d^n}{dt^n} L_{k,t}[f(x)] = \phi^{(n)}(t).$$

We have at once

$$\frac{\partial^n}{\partial t^n}\frac{t^{k-1}u^k}{(t+u)^{2k}} = (-1)^n \frac{\partial^n}{\partial u^n}\frac{t^{k-n-1}u^{k+n}}{(t+u)^{2k}},$$

so that

$$\frac{d^n}{dt^n}L_{k,t}[f(x)\ ] = (-1)^n d_k \int_0^\infty \phi(u)\frac{\partial^n}{\partial u^n}\frac{t^{k-n-1}u^{k+n}}{(t+u)^{2k}}\,du.$$

Conditions (7.2) now enable us to integrate by parts and obtain

$$\frac{d^n}{dt^n}L_{k,t}[f(x)] = \frac{1}{t^n}\int_0^\infty u^n\phi^{(n)}(u)\frac{t^{k-1}}{(t+u)^{2k}}\,du,$$

for values of k sufficiently large. But the asymptotic behavior of this integral as k becomes infinite was determined in §4. It clearly approaches the desired value.

Corollary 7.31. *Under the conditions of Corollary* 7.21

$$\lim_{k\to\infty}\frac{d^{n-1}}{dt^{n-1}}L_{k,t}[f(x)] = \alpha^{(n)}(t).$$

8. **A generalized Stieltjes transform.** The results of the preceding section suggest a way of inverting the general Stieltjes transform

$$F(x) = \int_0^\infty \frac{d\alpha(t)}{(x+t)^\rho},$$

where ρ is any positive number. We first prove

Lemma 8.11. *If* $t>0$, $x>0$, $\rho>0$, *then*

$$c_k\int_0^x \frac{(x-u)^{\rho-1}}{\Gamma(\rho)}(-u)^{k-1}\frac{\partial^{2k-1}}{\partial u^{2k-1}}\left[\frac{u^k\Gamma(\rho+1)}{(u+t)^{\rho+1}}\right]du = d_k\frac{t^{k-\rho}x^{k+\rho-1}}{(x+t)^{2k}}.$$

If $0\leqq u<t$,

$$\frac{(x-u)^{\rho-1}}{\Gamma(\rho)}(-u)^{k-1}\frac{\partial^{2k-1}}{\partial u^{2k-1}}\left[\frac{u^k\Gamma(\rho+1)}{(u+t)^{\rho+1}}\right]$$
$$= \sum_{n=k-1}^{\infty}\rho\binom{n+\rho}{n}\frac{(n+k)!}{(n-k+1)!}\frac{(-1)^{n+k-1}}{t^{n+\rho+1}}(x-u)^{\rho-1}u^n.$$

If $x<t$ integration term by term is permissible, so that

$$\int_0^x \frac{(x-u)^{\rho-1}}{\Gamma(\rho)}(-u)^{k-1}\frac{\partial^{2k-1}}{\partial u^{2k-1}}\left[\frac{u^k\Gamma(\rho+1)}{(u+t)^{\rho+1}}\right]du$$

$$= \sum_{n=k-1}^{\infty} \binom{n+\rho}{n} \frac{(n+k)!}{(n-k+1)!} \frac{(-1)^{n+k-1}}{t^{n+\rho+1}} \frac{x^{n+\rho}n!}{\Gamma(n+\rho+1)}$$

$$= (2k-1)! \frac{x^{\rho+k-1}}{t^{\rho+k}} \sum_{n=0}^{\infty} \binom{n+2k-1}{n} \left(\frac{-x}{t}\right)^n$$

$$= (2k-1)! \frac{t^{k-\rho}x^{k+\rho-1}}{(x+t)^{2k}} \cdot$$

It can now be seen by analytic continuation that the formula holds for all positive x and t.

LEMMA 8.12. *If the integral*

$$\int_0^\infty \frac{d\alpha(t)}{(x+t)^\rho} \qquad (\rho > 0)$$

converges, then

$$\alpha(t) = o(t^\rho) \qquad (t \to \infty). \tag{8.1}$$

The proof is similar to that of Theorem 1.2.

By use of these results we now prove

THEOREM 8.1. *If* $0<\rho<1$, *and if the integral*

$$F(x) = \int_0^\infty \frac{d\alpha(t)}{(x+t)^\rho}, \tag{8.2}$$

converges, then

$$\lim_{k\to\infty} \int_0^t (t-u)^{\rho-1} L_{k,u}[F(x)]du = \alpha(t). \tag{8.3}$$

Using Lemma 8.12 we have

$$F(x) = \rho \int_0^\infty \frac{\alpha(t)}{(x+t)^{\rho+1}} dt.$$

Then

$$L_{k,t}[F(x)] = c_k\rho \int_0^\infty \alpha(u)(-t)^{k-1} \frac{\partial^{2k-1}}{\partial t^{2k-1}} \left[\frac{t^k}{(t+u)^{\rho+1}}\right] du.$$

By uniform convergence we have for $0<\delta<y/2$

$$\int_\delta^{y-\delta} \frac{(y-t)^{\rho-1}}{\Gamma(\rho)} L_{k,t}[F(x)]dt$$

$$= c_k\rho \int_0^\infty \alpha(u)du \int_\delta^{y-\delta} \frac{(y-t)^{\rho-1}}{\Gamma(\rho)} (-t)^{k-1} \frac{\partial^{2k-1}}{\partial t^{2k-1}} \left[\frac{t^k}{(t+u)^{\rho+1}}\right] dt.$$

We may now replace δ by zero on both sides of this equation. To justify this step it is sufficient to show that

$$\int_0^y (y-t)^{\rho-1}dt \int_0^\infty |\alpha(u)|\, t^{k-1} \left| \frac{\partial^{2k-1}}{\partial t^{2k-1}} \left[\frac{t^k}{(t+u)^{\rho+1}} \right] \right| du$$

converges. For this it is sufficient to show that the integrals

$$(8.4) \qquad \int_0^y (y-t)^{\rho-1}dt \int_0^\infty |\alpha(u)| \frac{t^{2k-p-1}}{(t+u)^{\rho+2k-p}} du \qquad (p = 0, 1, \cdots, k)$$

all converge. But

$$\alpha(u) = \begin{cases} O(1) & (u \to 0), \\ O(u^\rho) & (u \to \infty), \end{cases}$$

so that

$$\int_0^\infty \frac{t^{2k-p-1}}{(t+u)^{\rho+2k-p}} |\alpha(u)|\, du = O(t^{-\rho}) \qquad (t \to 0;\ p = 0, 1, \cdots, k).$$

This proves the convergence of the integrals (8.4) and hence that

$$\int_0^y \frac{(y-t)^{\rho-1}}{\Gamma(\rho)} L_{k,t}[F(x)]dt$$

$$= c_k\rho \int_0^\infty \alpha(u)du \int_0^y \frac{(y-t)^{\rho-1}}{\Gamma(\rho)} (-t)^{k-1} \frac{\partial^{2k-1}}{\partial t^{2k-1}} \left[\frac{t^k}{(t+u)^{\rho+1}} \right] dt$$

$$= \frac{d_k}{\Gamma(\rho)} \int_0^\infty \alpha(u) \frac{u^{k-\rho} y^{k+\rho-1}}{(y+u)^{2k}} du.$$

If we use the known asymptotic expression of this last integral ($k \to \infty$), our theorem is established.

We illustrate the theorem by an example. Take

$$\begin{cases} \alpha(t) = 1 & (t > 0), \\ \alpha(0) = 0, \end{cases}$$

$$F(x) = \frac{1}{x^\rho} \cdot$$

Then

$$L_{k,u}[F(x)] = \frac{(k-\rho)(k-\rho-1)\cdots(1-\rho)\rho(\rho+1)\cdots(\rho+k-2)}{k!(k-2)!} \frac{1}{u^\rho},$$

$$\int_0^t (t-u)^{\rho-1} L_{k,u}[F(x)]du = \frac{\Gamma(\rho+k-1)\Gamma(\rho-k+1)}{k!(k-2)!} \qquad (t > 0),$$

and the right-hand side clearly approaches unity as k becomes infinite. This example shows in particular that the restriction $\rho<1$ in Theorem 8.1 was essential, for if $\rho>1$, the left-hand side of (8.3) need not converge.

If ρ is not less than unity we must proceed differently. In order to treat this case we introduce a new operator.

DEFINITION 8.1. *An operator $L^{\rho}_{k,t}[f(x)]$ is defined by the equation*

$$L^{\rho}_{k,t}[f(x)] = (-1)^k\Gamma(\rho)c_k[t^{2k-1}\{f^{(k)}(t)\}^{(-\rho)}]^{(k)}.$$

In this definition ρ is any positive number, k is a positive integer greater than ρ. The notation is understood to mean

$$\{f^{(k)}(t)\}^{(-\rho)} = \int_t^\infty \frac{(u-t)^{\rho-1}}{\Gamma(\rho)} f^{(k)}(u)du$$

if ρ is not an integer. If ρ is an integer

$$\{f^{(k)}(t)\}^{(-\rho)} = (-1)^\rho f^{(k-\rho)}(t).$$

It must not be supposed that for fractional ρ the function $\{f^{(k)}(t)\}^{(-\rho)}$ is the fractional derivative of $f(t)$ of order $k-\rho$, defined for $0<\rho<1$ by

$$\frac{d^k}{dt^k}\int_t^\infty \frac{(u-t)^{\rho-1}}{\Gamma(\rho)} f(u)du.$$

For, this integral need not exist in the present case. For example, if $f(t)=t^{-1/2}$, the integral does not exist if $\rho=1/2$. Yet the operator $L^{1/2}_{k,t}[f(x)]$ clearly exists for all integers k not less than unity.

We prove next

THEOREM 8.2. *If the integral* (8.2) *converges, then*

$$\lim_{k\to\infty}\int_0^t L^{\rho}_{k,u}[F(x)]du = \alpha(t) - \alpha(0+) \qquad (t>0).$$

We first prove that for k sufficiently large

$$\{f^{(k)}(x)\}^{(-\rho)} = (-1)^k\frac{(k-1)!}{\Gamma(\rho)}\int_0^\infty \frac{d\alpha(t)}{(x+t)^k}. \tag{8.5}$$

To see this we have

$$\begin{aligned}(-1)^kF^{(k)}(x) &= \frac{\Gamma(\rho+k)}{\Gamma(\rho)}\int_0^\infty \frac{d\alpha(t)}{(x+t)^{\rho+k}}\\ &= \frac{\Gamma(\rho+k+1)}{\Gamma(\rho)}\int_0^\infty \frac{\alpha(t)}{(x+t)^{\rho+k+1}}dt,\end{aligned} \tag{8.6}$$

$$\int_x^\infty \frac{(t-x)^{\rho-1}}{\Gamma(\rho)} F^{(k)}(t)dt$$

$$= (-1)^k \frac{\Gamma(\rho+k+1)}{\Gamma(\rho)} \int_x^\infty \frac{(t-x)^{\rho-1}}{\Gamma(\rho)} dt \int_0^\infty \frac{\alpha(u)}{(t+u)^{\rho+k+1}} du.$$

If $\delta>0$, $R>0$, $x+\delta<R$, the uniform convergence of the integral (8.6) shows us that

$$\int_x^\infty \frac{(t-x)^{\rho-1}}{\Gamma(\rho)} F^{(k)}(t)dt$$

$$= \lim_{\substack{\delta\to 0\\ R\to\infty}} (-1)^k \frac{\Gamma(\rho+k+1)}{\Gamma(\rho)} \int_0^\infty \frac{\alpha(u)}{\Gamma(\rho)} du \int_{x+\delta}^R \frac{(t-x)^{\rho-1}}{(t+u)^{\rho+k+1}} dt$$

$$= (-1)^k \frac{\Gamma(\rho+k+1)}{\Gamma(\rho)} \int_0^\infty \frac{\alpha(u)}{\Gamma(\rho)} du \int_x^\infty \frac{(t-x)^{\rho-1}}{(t+u)^{\rho+k+1}} dt$$

$$= (-1)^k \frac{\Gamma(k+1)}{\Gamma(\rho)} \int_0^\infty \frac{\alpha(u)}{(x+u)^{k+1}} du \tag{8.7}$$

provided the integral (8.7) converges absolutely. This it clearly does for $k>\rho$ by virtue of the relation (8.1). An integration by parts now gives (8.5). Then we obtain

$$\int_0^t L^\rho_{k,u}[F(x)]du = d_k t^k \int_0^\infty \frac{y^{k-1}\alpha(y)}{(t+y)^{2k}} dy - \frac{k-1}{k}\alpha(0+)$$

precisely as in the proof of Theorem 3.1. The theorem is now established by use of Corollary 3.13.

In a similar way we may obtain a generalization of Theorem 4.1.

THEOREM 8.3. *If $\phi(t)$ is integrable in* $(0, R)$ *for every positive* R, *and if* $F(x)$ *is defined by the convergent integral*

$$F(x) = \int_0^\infty \frac{\phi(t)}{(x+t)^\rho} dt \qquad (\rho>0),$$

then

$$\lim_{k\to\infty} L^\rho_{k,t}[F(x)] = \phi(t)$$

at every point t of the Lebesgue set for $\phi(u)$.

Let us illustrate these theorems by use of the same example as we used for illustration of Theorem 8.1.

$$F(x) = \frac{1}{x^\rho} = \int_0^\infty \frac{d\alpha(t)}{(x+t)^\rho} = \int_0^\infty \frac{\rho dt}{(x+t)^{\rho+1}} = \int_0^\infty \frac{\rho(\rho+1)t}{(x+t)^{\rho+2}}\,dt.$$

Direct application of Definition 8.1 gives us

$$L_{k,t}^{\rho}[F(x)] = 0,$$
$$L_{k,t}^{\rho+1}[F(x)] = \rho,$$
$$L_{k,t}^{\rho+2}[F(x)] = \rho(\rho+1)\frac{k+1}{k-2}\,t.$$

In each case the appropriate limit process gives the result predicted by Theorems 8.2 and 8.3.

9. **Uniqueness.** Of fundamental importance in later work will be the uniqueness theorem for the operator $L_{k,t}$. As a preliminary result we establish

THEOREM 9.1. *Let $f(t)$, $g(t)$ be functions of class C^{2k} in $0<t<\infty$, and let*

$$\lim\,[t^{2k-1}f^{(k)}(t)]^{(k-p)}g^{(p-1)}(t) = 0 \quad (t\to 0,\ t\to\infty;\ p = 1, 2, \cdots, k), \tag{9.1}$$

$$\lim\,[t^{2k-1}g^{(k)}(t)]^{(k-p)}f^{(p-1)}(t) = 0 \quad (t\to 0,\ t\to\infty;\ p = 1, 2, \cdots, k). \tag{9.2}$$

Then

$$\int_0^\infty [t^{2k-1}f^{(k)}(t)]^{(k)}g(t)dt = \int_0^\infty [t^{2k-1}g^{(k)}(t)]^{(k)}f(t)dt,$$

if either integral exists.

To verify this one has only to integrate successively by parts. Conditions (9.1) and (9.2) guarantee that at each stage the integrated part vanishes.

THEOREM 9.2. *If $f(x)$ is of class C^{2k-1} in the interval $0<x<\infty$, and if*

$$\lim\,[t^{2k-1}f^{(k-1)}(t)]^{(k-p)}(t+a)^{-p} = 0 \qquad (t\to 0,\ t\to\infty;\ p = 1, 2, \cdots, k;\ a>0), \tag{9.3}$$

$$\lim\,[t^{2k-1}(t+a)^{-k-1}]^{(k-p-2)}f^{(p)}(t) = 0 \qquad (t\to 0,\ t\to\infty;\ p = 0, 1, \cdots, k-2), \tag{9.4}$$

$$\begin{aligned} f(t) &= O(t^\mu) &&(t\to\infty,\ \mu<k-1),\\ f(t) &= O(t^{-\nu}) &&(t\to 0,\ \ \nu<k+1), \end{aligned} \tag{9.5}$$

then

$$\int_0^\infty \frac{L_{k,t}[f(x)]}{t+a}\,dt = \int_0^\infty L_{k,a}\left[\frac{1}{x+t}\right]f(t)dt.$$

This result may be obtained at once from Theorem 9.1. For in that theorem replace $g(t)$ by $(t+a)^{-1}$ and $f(t)$ by

$$f_1(t) = \int_1^t f(u)du.$$

We obtain

$$(9.6) \qquad \int_0^\infty \frac{L_{k,t}[f(x)]}{t+a}\,dt = (-1)^{2k-1}c_k \int_0^\infty \left[\frac{t^{2k-1}k!}{(t+a)^{k+1}}\right]^{(k)} f_1(t)dt.$$

Integrating the right-hand member by parts gives

$$(9.7) \qquad \begin{aligned} \int_0^\infty \frac{L_{k,t}[f(x)]}{t+a}\,dt &= c_k \int_0^\infty \left[\frac{t^{2k-1}k!}{(t+a)^{k+1}}\right]^{(k-1)} f(t)dt, \\ &= d_k \int_0^\infty \frac{a^{k-1}t^k}{(t+a)^{2k}} f(t)dt, \\ &= \int_0^\infty L_{k,a}\left[\frac{1}{x+t}\right] f(t)dt. \end{aligned}$$

It remains only to verify that the integral (9.7) converges. It clearly does by virtue of (9.5).

We come now to the uniqueness theorem.

THEOREM 9.3. *If $f(x)$ is of class C^∞ in $(0<x<\infty)$, if (9.3), (9.4) hold for $k=2, 3, 4, \cdots$, and if (9.5) holds for some positive μ and ν which are independent of k, then*

$$(9.8) \qquad f(a) = \lim_{k\to\infty} \int_0^\infty \frac{L_{k,t}[f(x)]}{t+a}\,dt \qquad (a>0).$$

The proof follows at once by use of (3.2).

Note that if $f_1(x)$ and $f_2(x)$ are two functions satisfying the condition of Theorem 9.3 and such that

$$L_{k,t}[f_1(x)] = L_{k,t}[f_2(x)]$$

for all sufficiently large integers k, then

$$f_1(x) = f_2(x) \qquad (0<x<\infty).$$

It is for this reason that the result may be regarded as a uniqueness theorem for the operator $L_{k,t}[f(x)]$.

10. **Sufficient conditions for uniqueness.** Conditions (9.3), (9.4), and (9.5) are sufficient for the application of Theorem 9.3. In this form, however, it may be difficult to determine, in any given case, whether a function satisfies

them or not. It is the purpose of the present section to replace them by conditions more easily applied.

THEOREM 10.1. *If*

$$
(10.1)\qquad
\begin{aligned}
f^{(k)}(x) &= o\left(\frac{1}{x^{k+1}}\right) && (x \to 0;\ k = 0, 1, 2, \cdots),\\
f^{(k)}(x) &= o\left(\frac{1}{x^{k}}\right) && (x \to \infty;\ k = 0, 1, 2, \cdots),
\end{aligned}
$$

then equation (9.8) *is true.*

For, if one expands (9.3) and (9.4) by Leibniz's rule, one sees that (9.3), (9.4), and (9.5) are true for all positive integers k by virtue of the relations (10.1). Note that no positive or negative integral power of x satisfies (10.1). For such functions $f(x)$ the right-hand side of (9.8) is zero.

For use in the proof of our next result we establish

LEMMA 10.21. *If k is a positive integer, and if*

$$f^{(k-1)}(t) = O\left(\frac{1}{t^k}\right) \qquad (t \to 0),$$

then

$$\int_0^t u^{2k+1}f^{(k)}(u)du = O(t^{k+1}) \qquad (t \to 0).$$

For, integration by parts gives for $\epsilon > 0$

$$\int_\epsilon^t u^{2k+1}f^{(k)}(u)du = t^{2k+1}f^{(k-1)}(t) - \epsilon^{2k+1}f^{(k-1)}(\epsilon) - (2k+1)\int_\epsilon^t u^{2k}f^{(k-1)}(u)du.$$

Making use of our hypothesis regarded $f^{(k-1)}(t)$, we obtain

$$
\begin{aligned}
\int_0^t u^{2k+1}f^{(k)}(u)du &= t^{2k+1}f^{(k-1)}(t) - (2k+1)\int_0^t u^{2k}f^{(k-1)}(u)du\\
&= O(t^{k+1}) \qquad (t \to 0).
\end{aligned}
$$

LEMMA 10.22. *If $f(x)$ is of class C^∞ in the interval $0 < x < \infty$, and if the limits*

$$\lim_{\epsilon \to 0+} \int_\epsilon^1 [u^{2k-1}f^{(k-1)}(u)]^{(k)}du \qquad (k = 1, 2, 3, \cdots)$$

exist, then

$$(-1)^k f^{(k)}(x) \sim \frac{Ak!}{x^{k+1}} \qquad (x \to 0+),$$

for a suitable constant A.

The hypothesis for $k=1$ assures us that the Cauchy value of the integral

$$\int_0^1 [uf(u)]'du$$

exists, and hence that there exists a constant A for which

$$tf(t) \sim A \qquad (t \to 0).$$

In particular

$$f(t) = O\left(\frac{1}{t}\right) \qquad (t \to 0).$$

We now proceed by induction and assume that

$$f^{(p)}(t) = O\left(\frac{1}{t^{p+1}}\right) \qquad (t \to 0;\ p = 0, 1, \cdots, k-1).$$

Since

$$\lim_{\epsilon \to 0+} \int_\epsilon^1 [u^{2k+1}f^{(k)}(u)]^{(k+1)}du$$

exists, it follows that

$$[u^{2k+1}f^{(k)}(u)]^{(k)} = O(1) \qquad (u \to 0).$$

Hence

$$\int_0^t [u^{2k+1}f^{(k)}(u)]^{(k)}du = O(t) \qquad (t \to 0),$$

$$[t^{2k+1}f^{(k)}(t)]^{(k-1)} - c_1 = O(t) \qquad (t \to 0),$$

for a suitable constant c_1. By successive integrations we have

$$\int_0^t u^{2k+1}f^{(k)}(u)du + P_k(t) = O(t^{k+1}) \qquad (t \to 0),$$

where $P_k(t)$ is a polynomial of degree at most equal to k. By Lemma 10.21

$$P_k(t) = O(t^{k+1}) \qquad (t \to 0).$$

But this is impossible unless $P_k(t)$ is identically zero. However,

$$\int_0^t [u^{2k+1}f^{(k)}(u)]'du + P_k'(t) = O(t^k),$$

$$t^{2k+1}f^{(k)}(t) = O(t^k),$$

$$f^{(k)}(t) = O\left(\frac{1}{t^{k+1}}\right) \qquad (t \to 0).$$

That is, this last relation must hold for all positive integers k. By use of a theorem of Hardy and Littlewood* the proof of the lemma is completed.

We can now prove

THEOREM 10.2. *If for each positive integer* k

$$\int_0^R L_{k,t}[f(x)]dt = O(R) \qquad (R\to\infty), \tag{10.2}$$

then

$$(-1)^k f^{(k)}(x) \sim \frac{Ak!}{x^{k+1}} \qquad (x\to 0;\ k=0,1,2,\cdots), \tag{10.3}$$

$$f^{(k)}(x) = O\left(\frac{1}{x^k}\right) \qquad (x\to\infty;\ k=0,1,2,\cdots), \tag{10.4}$$

where

$$A = \lim_{x\to 0+} xf(x). \tag{10.5}$$

The conclusions (10.3) and (10.5) follow at once from Lemma 10.22. To prove (10.4) we have

$$\int_0^R [t^{2k-1}f^{(k-1)}(t)]^{(k)}dt = O(R) \qquad (R\to\infty),$$

for each positive integer k. This shows that

$$[t^{2k-1}f^{(k-1)}(t)]^{(k-1)} = O(t) \qquad (t\to\infty),$$

$$t^{2k-1}f^{(k-1)}(t) = O(t^k) \qquad (t\to\infty),$$

from which (10.4) follows at once.

Our next result is

THEOREM 10.3. *If* $f(x)$ *satisfies* (10.2) *for each positive integer* k, *and if*

$$\lim_{x\to\infty} f(x) = 0, \tag{10.6}$$

then

$$f(a) = \lim_{k\to\infty}\int_0^\infty \frac{L_{k,t}[f(x)]}{t+a}dt + \frac{A}{a} \qquad (a>0), \tag{10.7}$$

where

$$A = \lim_{x\to 0+} xf(x).$$

---

* See, for example, E. Landau, *Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie*, Berlin, 1929, p. 58.

For, set

$$g(x) = f(x) - \frac{A}{x}\cdot$$

Then

$$g^{(k)}(x) = o\left(\frac{1}{x^{k+1}}\right) \qquad (x \to 0;\ k = 0, 1, 2, \cdots),$$

by Theorem 10.2. The same theorem shows that (10.4) is true for $g(x)$. This combined with (10.6) implies

$$g^{(k)}(x) = o\left(\frac{1}{x^{k}}\right) \qquad (x \to \infty;\ k = 0, 1, 2, \cdots).$$

Hence, by Theorem 10.1,

$$g(a) = \lim_{k\to\infty} \int_0^\infty \frac{L_{k,t}[g(x)]}{a+t}\,dt \qquad (a > 0).$$

Since

$$L_{k,t}[f(x)] = L_{k,t}[g(x)],$$

we see that (10.7) follows at once.

Corollary 10.31. *If the functions* $L_{k,t}[f(x)]$, $(k=1, 2, \cdots)$, *are all bounded, then* (10.6) *implies* (10.7).

Corollary 10.32. *If*

$$\int_0^\infty |\, L_{k,t}[f(x)]\,|^p dt < \infty \qquad (k = 1, 2, \cdots;\ p \geqq 1), \tag{10.8}$$

then (10.7) *holds.*

For, if $p=1$, then

$$\int_0^\infty |\, [tf(t)]'\,|\, dt < \infty,$$

so that (10.6) must hold. Clearly (10.8) implies (10.2) in this case. If $p>1$, Hölder's inequality gives

$$\int_0^R |\, L_{k,t}[f(x)]\,|\, dt \leqq \left[\int_0^\infty |\, L_{k,t}[f(x)]\,|^p dt\right]^{1/p} R^{(p-1)/p}, \tag{10.9}$$

so that (10.2) is satisfied. For $k=1$, (10.9) becomes

$$f(t) = O\left(\frac{1}{t^{1/p}}\right) \qquad (t \to \infty).$$

Hence (10.6) is satisfied. That is, Theorem 10.3 is applicable.

11. $\alpha(t)$ **of bounded variation.** Here we develop a necessary and sufficient condition that the equation (1.1) should have a solution $\alpha(t)$ of bounded variation in the infinite interval $(0, \infty)$.

THEOREM 11.1. *A necessary and sufficient condition that $f(x)$ should have the representation* (1.1) *with $\alpha(t)$ of bounded variation in* $(0, \infty)$ *is that*

$$\int_0^\infty | L_{k,t}[f(x)] | \, dt < M \qquad (k = 1, 2, \cdots), \tag{11.1}$$

where M is some constant independent of k.

To prove the necessity we have

$$L_{k,t}[f(x)] = d_k \int_0^\infty \frac{u^k t^{k-1}}{(t+u)^{2k}} \, d\alpha(u),$$

where

$$\int_0^\infty | d\alpha(u) | < \infty.$$

Then

$$\begin{aligned}
\int_0^\infty | L_{k,t}[f(x)] | \, dt &\leqq d_k \int_0^\infty t^{k-1} dt \int_0^\infty \frac{u^k}{(t+u)^{2k}} | d\alpha(u) | \\
&= d_k \int_0^\infty u^k | d\alpha(u) | \int_0^\infty \frac{t^{k-1}}{(t+u)^{2k}} dt \\
&= \frac{k-1}{k} \int_0^\infty | d\alpha(u) | \leqq \int_0^\infty | d\alpha(u) | .
\end{aligned}$$

This proves the necessity.

For the sufficiency, we have by Corollary 10.32

$$f(a) = \lim_{k\to\infty} \int_0^\infty \frac{L_{k,t}[f(x)]}{t+a} dt + \frac{A}{a} \qquad (a > 0),$$

$$A = \lim_{x\to 0+} x f(x).$$

By a theorem of Helly* we can pick from the set of functions

$$\alpha_k(t) = \int_0^t L_{k,u}[f(x)] du \qquad (k = 1, 2, \cdots)$$

* E. Helly, *Über lineare Funktionaloperationen*, Wiener Sitzungsberichte, vol. 121 (1921), p. 265.

a subset $\alpha_{k_i}(t)$ which approaches a function $\alpha^*(t)$ of bounded variation in the interval $(0, \infty)$. Then

$$f(a) = \lim_{i\to\infty} \int_0^\infty \frac{d\alpha_{k_i}(t)}{t+a} + \frac{A}{a} \qquad (a > 0).$$

By the Helly-Bray† theorem we may take the limit under the integral sign and obtain

$$\begin{aligned} f(a) &= \int_0^\infty \frac{d\alpha^*(t)}{t+a} + \frac{A}{a} \\ &= \int_0^\infty \frac{d\alpha(t)}{t+a} \qquad (a > 0), \end{aligned}$$

where $\alpha(t)$ vanishes at the origin and differs from $\alpha^*(t)$ by the constant A for positive values of t. This completes the proof of the theorem.

12. $\alpha(t)$ **non-decreasing.** Let us introduce

DEFINITION 12.1. *A function $f(x)$ satisfies Property* A *if and only if*

$$L_{k,t}[f(x)] \geqq 0 \qquad (t > 0;\ k = 0, 1, 2, \cdots).$$

Clearly this is equivalent to

$$f(x) \geqq 0, \quad (-1)^{k-1}[x^{2k-1}f^{(k-1)}(x)]^{(k)} \geqq 0 \qquad (x > 0;\ k = 1, 2, \cdots),$$

or to

$$f(x) \geqq 0, \qquad (-1)^{k-1}[x^k f(x)]^{(2k-1)} \geqq 0 \qquad (x > 0;\ k = 1, 2, \cdots).$$

In the proof of our result we shall need

LEMMA 12.11. *If $\phi(x)$ is of class C^1 in $0<x\leqq 1$ and if $\phi'(x)$ is bounded on one side in that interval, then $-\phi(x)$ is bounded on the same side there.*

The proof is obvious.

THEOREM 12.1. *If $f(x)$ has Property* A, *then the relations* (10.3) *and* (10.5) *hold.*

Since $xf(x)$ is a positive increasing function, it follows that the constant A of (10.5) is well defined. It will be sufficient to prove that

$$f^{(p)}(x) = O\left(\frac{1}{x^{p+1}}\right) \qquad (x \to 0;\ p = 0, 1, 2, \cdots).$$

Since this has been proved for $k=0$, we may proceed by induction. Let us

† See, for example, G. C. Evans, *The Logarithmic Potential, Discontinuous Dirichlet and Neumann Problems*, American Mathematical Society Colloquium Publications, vol. 6, 1927, p. 15.

grant then that these relations hold for $p=0, 1, 2, \cdots, k-2$. By hypothesis

$$(-1)^{k-1}[x^k f(x)]^{(2k-1)} \geqq 0 \qquad (x > 0).$$

By Lemma 12.11

$$[x^k f(x)]^{(k-1)} < M \qquad (0 < x \leqq 1)$$

for a suitable constant M. Also

$$(-1)^{k-2}[x^{k-1} f(x)]^{(2k-3)} \geqq 0$$

from which we deduce in the same way that

$$[x^{k-1} f(x)]^{(k-1)} > N \qquad (0 < x \leqq 1)$$

for a suitable constant N. But

$$[x^k f(x)]^{(k-1)} = x[x^{k-1} f(x)]^{(k-1)} + (k-1)[x^{k-1} f(x)]^{(k-2)}.$$

Since the second term on the right-hand side is $O(1)$ by our assumption (10.6), it follows that

$$Nx + O(1) < [x^k f(x)]^{(k-1)} < M \qquad (0 < x \leqq 1).$$

Hence

$$[x^k f(x)]^{(k-1)} = O(1) \qquad (x \to 0).$$

Expanding by Leibniz's rule we see that

$$f^{(k-1)}(x) = O\left(\frac{1}{x^k}\right),$$

so that the induction is complete. Hence (10.3) is established.

THEOREM 12.2. *If $f(x)$ has Property* A, *then there exist constants $A_0, A_1, \cdots$ such that*

$$(12.1) \qquad (-1)^k f^{(k)}(x) \geqq \frac{A_k}{x^{k+1}} \qquad (k = 0, 1, 2, \cdots; 0 < x < \infty).$$

Since

$$(-1)^{k-1}[x^{2k-1} f^{(k-1)}(x)]^{(k)} \geqq 0,$$

it follows by successive integration that

$$(-1)^{k-1}[x^{2k-1} f^{(k-1)}(x)] \geqq A_{k-1} x^{k-1} \qquad (1 \leqq x < \infty).$$

By Theorem 12.1 a similar inequality holds in the interval $(0 < x \leqq 1)$, so that (12.1) follows at once.

THEOREM 12.3. *If $f(x)$ has Property* A *and if $xf(x)$ approaches a limit as k becomes infinite, then*

$$\int_0^\infty L_{k,t}[f(x)]dt = \frac{k-1}{k}\left[\lim_{x\to\infty} xf(x) - \lim_{x\to 0} xf(x)\right].$$

For, if

$$f(x) \sim \frac{B}{x} \qquad (x \to \infty),$$

then Theorem 12.2 with the addition of the Hardy-Littlewood result referred to earlier shows us that

$$(12.2) \qquad (-1)^k f^{(k)}(x) \sim \frac{Bk!}{x^{k+1}} \qquad (x \to \infty).$$

It is easily seen that the relations (12.2) imply that

$$\lim_{t\to\infty} (-1)^{k-1}[t^{2k-1}f^{(k-1)}(t)]^{(k-1)} = (k-1)!(k-1)!B.$$

There is of course a similar result for $t \to 0$ following from the relations (10.3) which are necessary consequences of Property A. Hence

$$\int_0^\infty L_{k,t}[f(x)]dt = c_k \int_0^\infty (-1)^{k-1}[t^{2k-1}f^{(k-1)}(t)]^{(k)}dt$$

$$= \frac{(k-1)!(k-1)!}{k!(k-2)!}[B - A].$$

This proves the theorem. It is to be noted that the existence of B added to our hypothesis in this theorem is not a consequence of Property A, as one sees by the examples $f(x) = 1$ and $f(x) = (x)^{-1/2}$. Both satisfy the property but in each case B fails to exist. We can now treat the case† of bounded non-decreasing functions $\alpha(t)$.

THEOREM 12.4. *A necessary and sufficient condition that $f(x)$ should have the form* (1.1) *with $\alpha(t)$ bounded non-decreasing is that $f(x)$ should have the Property* A *and that $xf(x)$ should approach a limit as x becomes infinite. Further,*

$$(12.3) \qquad \alpha(\infty) - \alpha(0+) = \lim_{k\to\infty} \int_0^\infty L_{k,t}[f(x)]dt.$$

The necessity of Property A follows from an inspection of the relations

† The author treated this case by another method in an earlier paper: D. V. Widder, *A classification of generating functions*, these Transactions, vol. 39 (1936), p. 244.

$$L_{0,t}[f(x)] = f(t),$$

$$L_{k,t}[f(x)] = d_k t^{k-1} \int_0^\infty \frac{u^k}{(t+u)^{2k}} d\alpha(u) \qquad (k = 1, 2, \cdots).$$

Moreover, it is obvious that

$$\lim_{x\to\infty} xf(x) = \lim_{x\to\infty} \int_0^\infty \frac{x d\alpha(t)}{x+t} = \alpha(\infty).$$

To prove the sufficiency we first appeal to Theorem 12.3. This shows that

$$\int_0^\infty L_{k,t}[f(x)]dt \leqq B - A \qquad (k = 1, 2, \cdots),$$

where A, B are defined as in the proof of Theorem 12.3. Hence (11.1) is satisfied, and, by Theorem 11.1, $f(x)$ has the representation (1.1) with $\alpha(t)$ of bounded variation in $(0, \infty)$. To show that $\alpha(t)$ is non-decreasing we now appeal to Theorem 3.1. Clearly, on the assumption of Property A, the functions

$$\int_0^t L_{k,u}[f(x)]du \qquad (k = 1, 2, \cdots)$$

are non-decreasing functions of t.

Finally, (12.3) is a direct result of Corollary 6.31.

We turn next to the case of unbounded non-decreasing functions $\alpha(t)$. For the discussion of this case we need

LEMMA 12.51. *If $f(x)$ satisfies Property* A, *then it approaches a limit as x becomes infinite.*

For, since

$$-[u^2f(u)]^{(3)} \geqq 0 \qquad (u > 0),$$

we have for $0<y<x$ by successive integrations

$$-x^2f(x) + y^2f(y) + (x-y)[y^2f(y)]' + \frac{(x-y)^2}{2}[y^2f(y)]'' \geqq 0,$$

whence

$$\limsup_{x\to\infty} f(x) = E \leqq \tfrac{1}{2}[y^2f(y)]'' \qquad (y > 0).$$

Successive integration of the inequality

$$[u^2f(u)]'' - 2E \geqq 0$$

gives

$$x^2f(x) - y^2f(y) - (x - y)[y^2f(y)]' - E(x - y)^2 \geqq 0 \qquad (0 < y < x).$$

Hence

$$\liminf_{x\to\infty} f(x) \geqq E,$$

or

$$\lim_{x\to\infty} f(x) = E \geqq 0. \tag{12.4}$$

LEMMA 12.52. *If $f(x)$ satisfies Property* A, *then for each non-negative integer k the function $[x^kf(x)]^{(k)}$ is completely monotonic for $x>0$,*

$$(-1)^n[x^kf(x)]^{(k+n)} \geqq 0 \qquad (n = 0, 1, 2, \cdots).$$

By (12.4) and (12.1) it follows from the Hardy-Littlewood result quoted earlier that

$$f^{(k)}(x) = o\left(\frac{1}{x^k}\right) \qquad (k = 1, 2, \cdots; x \to \infty).$$

Since

$$(-1)^{k-1}[x^kf(x)]^{(2k-1)} \geqq 0 \qquad (x > 0),$$

and since

$$\lim_{x\to\infty} [x^kf(x)]^{(n)} = 0 \quad (n = k+1, k+2, \cdots),$$

it follows by successive integrations to infinity that

$$(-1)^n[x^kf(x)]^{(k+n)} \geqq 0 \qquad (n = 1, 2, \cdots, k-1; x > 0), \tag{12.5}$$

$$[x^kf(x)]^{(k)} \geqq E \geqq 0. \tag{12.6}$$

It remains to show that (12.5) holds for $n = k, k+1, \cdots$.

Let r be a positive integer, and replace $f(x)$ by $x^rf(x)$ in Lemma 3.11. We obtain

$$[x^{2k-1}\{x^rf(x)\}^{(k-1)}]^{(k)} = x^{k-1}[x^{k+r}f(x)]^{(2k-1)}.$$

Hence by (12.5) and (12.6)

$$(-1)^{k-r-1}[x^{k+r}f(x)]^{(2k-1)} \geqq 0 \qquad (k \geqq r+1),$$

$$(-1)^{k-r-1}[x^{2k-1}\{x^rf(x)\}^{(k-1)}]^{(k)} \geqq 0 \qquad (k \geqq r+1). \tag{12.7}$$

But

$$\lim_{x\to 0+} [x^{2k-1}\{x^rf(x)\}^{(k-1)}]^{(p)} = 0 \qquad (p = 0, 1, 2, \cdots, k-1; r > 0),$$

$$\lim_{x\to 0+} (-1)^{k-1}[x^{2k-1}f^{(k-1)}(x)]^{(k-1)} = (k-1)!(k-1)!\,A \geqq 0.$$

Hence successive integrations of (12.7) from zero give

$$(-1)^{k-r-1}[x^r f(x)]^{(k-1)} \geqq 0 \qquad (k \geqq r+1),$$

and this completes the proof of the lemma.

LEMMA 12.53. *If $f(x)$ has Property* A, *and if $\delta>0$, then $f(x+\delta)$ has the same property.*

For,

$$(12.8)\qquad (-1)^{k-1}[x^k f(x+\delta)]^{(2k-1)} = \sum_{p=0}^{k}\binom{k}{p}(-1)^{p+k-1}[(x+\delta)^{k-p}f(x+\delta)]^{(2k-1)}.$$

By Lemma 12.52

$$(-1)^{k+p-1}[(x+\delta)^{k-p}f(x+\delta)]^{(2k-1)} \geqq 0 \qquad (p = 0, 1, \cdots, k),$$

so that every term in the sum (12.8) is non-negative.

LEMMA 12.54. *If $f(x)$ has Property* A *and if $\delta>0$, then*

$$F(x) = \frac{f(x+\delta)-f(\delta)}{-x}$$

has the property and

$$(12.9)\qquad \lim_{x\to\infty} xF(x) = f(\delta) - f(\infty).$$

For

$$(-1)^{k-1}[x^k F(x)]^{(2k-1)} = (-1)^k[x^{k-1}f(x+\delta)]^{(2k-1)} \qquad (k = 1, 2, \cdots).$$

But the right-hand side is non-negative by Lemmas 12.52 and 12.53. Also

$$F(x) = \frac{f(x+\delta)-f(\delta)}{-x} = -f'(\xi) \geqq 0 \qquad (\delta < \xi < \delta + x),$$

so that $F(x)$ has Property A. By Lemma 12.51 we deduce (12.9).

By use of this last result we can now prove the main result of the section.

THEOREM 12.5. *Property* A *for the function $f(x)$ is necessary and sufficient that it should have the form*

$$(12.10)\qquad f(x) = E + \int_0^\infty \frac{d\alpha(t)}{x+t},$$

where $\alpha(t)$ is non-decreasing and $E\geqq 0$.

The proof of the necessity is made as in Theorem 12.4.

For the sufficiency we have at once by Lemma 12.54 and Theorem 12.4 for a given positive δ

$$F(x) = \frac{f(x+\delta) - f(\delta)}{-x} = \int_0^\infty \frac{d\beta(t)}{x+t},$$

where $\beta(t)$ is non-decreasing and bounded. In fact

$$\beta(0) = 0, \qquad \beta(\infty) = f(\delta) - f(\infty).$$

But $\beta(t)$ must be constantly zero in $(0, \delta)$, for otherwise it would have a point of increase there and by Theorem 1.7, $f(x)$ would have a singularity for some positive x. But, by Lemma 12.52, $f(x)$ is completely monotonic and hence analytic for $x>0$. Hence

$$\begin{aligned} f(x+\delta) &= f(\delta) - \int_0^\infty \frac{x}{x+t}\,d\beta(t) = f(\infty) + \int_0^\infty d\beta(t) - \int_0^\infty \frac{x}{x+t}\,d\beta(t) \\ &= f(\infty) + \int_\delta^\infty \frac{t}{x+t}\,d\beta(t) = f(\infty) + \int_0^\infty \frac{t+\delta}{x+t+\delta}\,d\beta(t+\delta), \\ f(x) &= f(\infty) + \int_0^\infty \frac{(t+\delta)}{x+t}\,d\beta(t+\delta) \end{aligned}$$

$$= f(\infty) + \int_0^\infty \frac{d\alpha(t)}{x+t} \qquad (x > \delta), \tag{12.11}$$

where

$$\alpha(t) = \int_0^t (u+\delta)d\beta(u+\delta).$$

Clearly $\alpha(t)$ is non-decreasing. It is independent of δ by Theorem 1.6. Hence (12.11) holds for all $x>0$, and our theorem is proved.

13. $\phi(t)$ **of class** L^p, $p>1$. In this section we deduce conditions on $f(x)$ which will insure its representation in the form (4.1) with

$$\int_0^\infty |\phi(t)|^p dt < \infty \tag{13.1}$$

for some constant $p>1$. The result is

THEOREM 13.1. *A necessary and sufficient condition that* $f(x)$ *should have the form* (4.1) *with* $\phi(t)$ *satisfying* (13.1) *is that*

$$\int_0^\infty |L_{k,t}[f(x)]|^p dt < M \qquad (k = 1, 2, \cdots), \tag{13.2}$$

$$\lim_{x\to 0+} xf(x) = 0, \tag{13.3}$$

where M is some constant independent of k.

For the necessity we have from §4

$$L_{k,t}[f(x)] = d_k t^{k-1}\int_0^\infty \frac{u^k}{(t+u)^{2k}}\phi(u)du \qquad (k = 1, 2, \cdots).$$

These integrals all converge absolutely since

$$\int_0^\infty \frac{u^k}{(t+u)^{2k}}\,|\,\phi(u)\,|\,du \leqq \left[\int_0^\infty |\,\phi(u)\,|^p du\right]^{1/p}\left[\frac{\Gamma(qk+1)\Gamma(qk-1)}{t^{qk-1}\Gamma(2qk)}\right]^{1/q},$$
$$\frac{1}{p}+\frac{1}{q} = 1,$$

by Hölder's inequality. We also have for $k>1$

$$|\,L_{k,t}[f(x)]\,|^p \leqq d_k t^{k-1}\int_0^\infty \frac{u^k}{(t+u)^{2k}}\,|\,\phi(u)\,|^p du\left[d_k\int_0^\infty \frac{t^{k-1}u^k}{(t+u)^{2k}}du\right]^{p/q},$$

(13.4) $$\int_0^\infty |\,L_{k,t}[f(x)]\,|^p dt \leqq d_k\int_0^\infty t^{k-1}dt\int_0^\infty \frac{u^k}{(t+u)^{2k}}\,|\,\phi(u)\,|^p du$$
$$= \frac{k-1}{k}\int_0^\infty |\,\phi(u)\,|^p du = \int_0^\infty |\,\phi(u)\,|^p du.$$

For $k=1$, this argument fails. However, in this case we can obtain our result by use of Hilbert's double-integral theorem.†

(13.5) $$L_{1,t}[f(x)] = \int_0^\infty \frac{u\phi(u)}{(t+u)^2}du,$$
$$\int_0^\infty |\,L_{1,t}[f(x)]\,|^p dt \leqq r^p\int_0^\infty |\,\phi(u)\,|^p du,$$

where

$$r = \frac{\Gamma\left(2-\frac{1}{p}\right)\Gamma\left(\frac{1}{p}\right)}{\Gamma(2)}.$$

Clearly (13.3) also holds, since

$$\lim_{x\to 0+} xf(x) = \lim_{t\to 0+}\int_0^t \phi(u)du.$$

† See G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities*, Cambridge, 1934, p. 229, Theorem 319.

Hence the necessity of (13.2) is established.

Conversely, we see that (13.2) implies, by Corollary 10.32, that

$$f(a) = \lim_{k\to\infty} \int_0^\infty \frac{L_{k,t}[f(x)]}{t+a}\,dt + \frac{A}{a} \qquad (a > 0),$$

$$A = \lim_{x\to 0+} xf(x).$$

Furthermore, (13.2) implies† the existence of a subset k_i of all the integers k and a function $\phi(t)$ of class L^p in $(0, \infty)$ such that

$$\lim_{i\to\infty} \int_0^\infty \frac{L_{k_i,t}[f(x)]}{t+a}\,dt = \int_0^\infty \frac{\phi(t)}{t+a}\,dt.$$

Hence

$$f(x) = \frac{A}{x} + \int_0^\infty \frac{\phi(t)}{x+t}\,dt \qquad (x > 0). \tag{13.6}$$

But A is zero by virtue of (13.3), so that the theorem is established.

COROLLARY 13.11. *Conditions* (13.2) *are necessary and sufficient that* $f(x)$ *should have the representation* (13.6) *with* $\phi(t)$ *satisfying* (13.1).

COROLLARY 13.12. *If* $f(x)$ *has the representation* (4.1), (13.1), *then*

$$\lim_{k\to\infty} \int_0^\infty |L_{k,t}[f(x)]|^p dt = \int_0^\infty |\phi(t)|^p dt.$$

For, Fatou's lemma gives

$$\int_0^\infty |\phi(t)|^p dt \leqq \liminf \int_0^\infty |L_{k,t}[f(x)]|^p dt.$$

This combined with (13.4) gives the result.

14. **Continuation,** $p=1$. That Theorem 13.1 can not hold for $p=1$ one sees from Theorem 11.1. For this case we prove

THEOREM 14.1. *A necessary and sufficient condition that* $f(x)$ *should have the form* (4.1) *with* $\phi(t)$ *of class L in* $(0, \infty)$ *is that the functions* $L_{k,t}[f(x)]$, $(k=1, 2, \cdots)$, *should all be of class L and that*

$$\lim_{k,l\to\infty} \int_0^\infty |L_{k,t}[f(x)] - L_{l,t}[f(x)]|\,dt = 0, \tag{14.1}$$

$$\lim_{x\to 0+} xf(x) = 0. \tag{14.2}$$

† See S. Banach, *Opérations Linéaires*, p. 130. The proof there given is easily extended to the case of an infinite interval.

If $f(x)$ has the form (4.1) with $\phi(t)$ of class L, then

$$\int_0^\infty | L_{k,t}[f(x)] | \, dt \leqq d_k \int_0^\infty dt \int_0^\infty \frac{t^{k-1}u^k}{(t+u)^{2k}} | \phi(u) | \, du$$

$$= \int_0^\infty u^k | \phi(u) | \, du \int_0^\infty \frac{t^{k-1}}{(t+u)^{2k}} \, dt \leqq \int_0^\infty | \phi(u) | \, du \qquad (k = 1, 2, \cdots),$$

so that the first part of our condition is necessary. For the second part we have

$$| L_{k,t}[f(x)] - \phi(t) | \leqq d_k \int_0^\infty \frac{t^{k-1}u^k}{(t+u)^{2k}} | \phi(u) - \phi(t) | \, du$$

$$= d_k \int_0^\infty \frac{u^k}{(u+1)^{2k}} | \phi(tu) - \phi(t) | \, du.$$

Hence

$$\int_0^\infty | L_{k,t}[f(x)] - \phi(t) | \, dt \leqq d_k \int_0^\infty \frac{u^k}{(u+1)^{2k}} g(u) du,$$

where

$$g(u) = \int_0^\infty | \phi(tu) - \phi(t) | \, dt.$$

But $g(1) = 0$, $g(t)$ is continuous at $u = 1$, and a constant M exists such that

$$| g(u) | < Mu^{-1} + M \qquad (0 < u < \infty).$$

Under these conditions

$$\lim_{k\to\infty} d_k \int_0^\infty \frac{u^k}{(u+1)^{2k}} g(u) du = g(1) = 0$$

by Corollary 4.11. From this (14.1) is immediate.

Conversely, the assumption (14.1) implies the existence of a function $\phi(t)$ of class L such that

$$\lim_{k\to\infty} \int_0^\infty | L_{k,t}[f(x)] - \phi(t) | \, dt = 0, \tag{14.3}$$

$$\lim_{k\to\infty} \int_0^\infty | L_{k,t}[f(x)] | \, dt = \int_0^\infty | \phi(t) | \, dt. \tag{14.4}$$

Equation (14.4) combined with (14.1) implies (11.1) for a suitable constant M. Hence by Theorem 11.1

$$f(x) = \int_0^\infty \frac{d\alpha(t)}{x+t},$$

where $\alpha(t)$ is of bounded variation in $(0, \infty)$. But

$$\left| \int_0^u L_{k,t}[f(x)]dt - \int_0^u \phi(t)dt \right| \leqq \int_0^\infty | L_{k,t}[f(x)] - \phi(t) | \, dt \qquad (u > 0).$$

Hence by (14.3)

$$\lim_{k\to\infty} \int_0^u L_{k,t}[f(x)]dt = \int_0^u \phi(t)dt.$$

But by Theorem 3.1

$$\lim_{k\to\infty} \int_0^u L_{k,t}[f(x)]dt = \alpha(u) - \alpha(0+) \qquad (u > 0).$$

By (14.2)

$$\lim_{x\to 0+} xf(x) = \alpha(0+) = 0,$$

so that

$$\alpha(u) = \int_0^u \phi(t)dt \qquad (u \geqq 0).$$

This completes the proof of the theorem.

COROLLARY 14.11. *If $f(x)$ has the form* (4.1) *with $\phi(t)$ of class L, then*

$$\lim_{k\to\infty} \int_0^\infty L_{k,t}[f(x)]dt = \int_0^\infty \phi(t)dt.$$

15. $\phi(t)$ **bounded.** To conjecture a condition for this case one would naturally allow p to become infinite in (13.2). This would lead to

$$| L_{k,t}[f(x)] | < N \qquad (k = 1, 2, \cdots). \tag{15.1}$$

But note that for $k=1$ we have (13.5) and that

$$\lim_{p\to\infty} r = \lim_{p\to\infty} \Gamma\left(2 - \frac{1}{p}\right)\Gamma\left(\frac{1}{p}\right) = \infty.$$

In fact (15.1) is not necessary for the boundedness of $\phi(t)$. For, let $\phi(t)$ be equal to unity in (0, 1) and zero elsewhere. Then

$$L_{1,t}[f(x)] = \log\left(1 + \frac{1}{t}\right) - \frac{1}{t+1} \qquad (t > 0),$$

and this function becomes infinite as t approaches zero. We may overcome this difficulty by replacing (15.1) ($k=1$) by a condition on $f(x)$ of a slightly different type. The result is stated in

THEOREM 15.1. *A necessary and sufficient condition that $f(x)$ can be represented in the form* (4.1) *with $\phi(t)$ bounded is that*

$$|L_{k,t}[f(x)]| < M \qquad (t > 0;\ k = 2, 3, \cdots), \tag{15.2}$$

$$\lim_{x\to 0} xf(x) = 0, \tag{15.3}$$

$$\lim_{x\to\infty} f(x) = 0 \tag{15.4}$$

for a suitable constant M.

If $f(x)$ has the form (4.1), and if

$$|\phi(t)| < M \qquad (0 < t < \infty),$$

then

$$|L_{k,t}[f(x)]| \leqq d_k \int_0^\infty \frac{u^k t^{k-1}}{(u+t)^{2k}} |\phi(u)|\, du \quad (k = 2, 3, \cdots),$$

so that (15.2) is satisfied. Also

$$\lim_{x\to 0+} \int_0^\infty \frac{x\phi(t)}{x+t} dt = \lim_{u\to 0+} \int_0^u \phi(t)dt = 0,$$

$$\lim_{x\to\infty} \int_0^\infty \frac{\phi(t)}{x+t} dt = 0,$$

so that (15.3) and (15.4) also hold.

Conversely, (15.2) implies (10.2) at least for $k=2, 3, \cdots$. Also (15.3) and (15.4) imply (10.2) for $k=1$. Hence (10.3) and (10.4) hold. But these combined with (15.3) and (15.4) give

$$f^{(k)}(x) = o\left(\frac{1}{x^{k+1}}\right) \qquad (x \to 0;\ k = 0, 1, 2, \cdots),$$

$$f^{(k)}(x) = o\left(\frac{1}{x^k}\right) \qquad (x \to \infty;\ k = 0, 1, 2, \cdots).$$

Hence we obtain by successive integration by parts the identity

$$c_k(-1)^{k-1}\int_0^\infty \frac{[t^{2k-1}f^{(k-1)}(t)]^{(k)}}{(x+t)^2} dt = -d_k x^{k-2}\int_0^\infty \frac{t^{k+1}}{(x+t)^{2k}} f'(t)dt.$$

By (3.2)

$$- f'(x) = \lim_{k\to\infty} \int_0^\infty \frac{L_{k,t}[f(x)]}{(x+t)^2}\, dt.$$

Furthermore (15.2) implies the existence† of a subset k_i of the integers k and a bounded function $\phi(t)$ such that

$$- f'(x) = \lim_{i\to\infty} \int_0^\infty \frac{L_{k_i,t}[f(x)]}{(x+t)^2}\, dt = \int_0^\infty \frac{\phi(t)}{(x+t)^2}\, dt.$$

Now let $0<x<y$. Since the integral

$$\int_0^\infty \frac{\phi(t)}{(x+t)^2}\, dt$$

is uniformly convergent in any closed interval of the positive x-axis, we have

$$f(x) - f(y) = (y - x) \int_0^\infty \frac{\phi(t)}{(x+t)(y+t)}\, dt.$$

Hence, for any fixed positive x, (15.4) gives

$$(15.5) \qquad \int_0^\infty \frac{\phi(t)}{(x+t)(y+t)}\, dt \sim \frac{f(x)}{y} \qquad (y \to \infty).$$

Moreover, since $\phi(t)$ is bounded, there exists a constant N such that

$$\frac{\phi(t)}{x+t} \geqq -\frac{N}{t} \qquad (0 < t < \infty).$$

Hence we are in a position to apply a Tauberian theorem of Hardy and Littlewood‡ to the integral (15.5) considered as a function of y. The conclusion is that

$$f(x) = \int_0^\infty \frac{\phi(t)}{x+t}\, dt,$$

which is what we set out to prove.

COROLLARY 15.1. *If $f(x)$ has the form* (4.1) *with $\phi(t)$ bounded, then*§

$$\lim_{k\to\infty} \underset{0<t<\infty}{\text{l.u.b.}} \left| L_{k,t}[f(x)] \right| = \underset{0<t<\infty}{\text{true max}} \left| \phi(t) \right|.$$

---

† See, for example, S. Banach, loc. cit., p. 130.

‡ G. H. Hardy and J. E. Littlewood, *Notes on the theory of series* (XI): *on Tauberian theorems*, Proceedings of the London Mathematical Society, vol. 30 (1930), p. 33.

§ For definition of true max see, for example, S. Banach, loc. cit., p. 227.

16. **A more general case.** We next investigate what functions $f(x)$ can be represented by a convergent integral of the form

$$f(x) = \int_0^\infty \frac{d\alpha(t)}{x+t},$$

with no restriction on $\alpha(t)$ except that it should be of bounded variation in every finite interval, and bounded in the infinite interval. To treat this case we need certain preliminary results which we now establish. We introduce a new operator $M_{k,t}[f(x)]$ by the

DEFINITION.

$$M_{k,t}[f(x)] = (-1)^{k-1}c_k[t^{2k-1}f^{(k-1)}(t)]^{(k-1)} \qquad (k = 2, 3, 4, \cdots),$$
$$M_{1,t}[f(x)] = tf(t).$$

Our first result is contained in

THEOREM 16.1. *If*

(16.1) $$f^{(n)}(t) = o\left(\frac{1}{t^{n+2}}\right) \qquad (t \to 0;\ n = 0, 1, 2, \cdots),$$

(16.2) $$f^{(n)}(t) = o\left(\frac{1}{t^n}\right) \qquad (t \to \infty;\ n = 0, 1, 2, \cdots),$$

then

$$f(x) = \lim_{k\to\infty} \int_0^\infty \frac{M_{k,t}[f(x)]}{(x+t)^2}\,dt \qquad (x > 0).$$

By an application of Theorem 9.1 or directly by integration by parts one shows that

$$\int_0^\infty \frac{M_{k,t}[f(x)]}{(x+t)^2}\,dt = d_k \int_0^\infty \frac{x^{k-1}t^k}{(x+t)^{2k}} f(t)dt.$$

Equation (3.2) now gives the result desired.

We shall next show that conditions (16.1), (16.2) follow automatically from the boundedness of $M_{k,t}[f(x)]$.

THEOREM 16.2. *If*

(16.3) $$M_{k,t}[f(x)] = O(1) \qquad (t \to \infty,\ t \to 0;\ k = 1, 2, \cdots),$$

then (16.1) *and* (16.2) *are true.*

For, one sees easily that (16.3) implies

$$f^{(k)}(t) = O\left(\frac{1}{t^{k+1}}\right) \qquad (t \to \infty,\ k = 0, 1, 2, \cdots),$$

of which (16.2) is a trivial consequence. Furthermore (16.3) implies

$$[t^{2k-1}f^{(k-1)}(t)]^{(k-1)} = O(1) \qquad (t \to 0), \tag{16.4}$$

$$[t^{2k-1}f^{(k-1)}(t)]^{(k-2)} = O(1) \qquad (t \to 0). \tag{16.5}$$

If we now assume (16.1) for $(n=0, 1, 2, \cdots, 2k-4)$ we see that it also holds for $n=2k-3$ by (16.5) and for $n=2k-2$ by (16.4). Since it obviously holds for $n=0$ by (16.4), $k=1$, it must hold in general.

By use of these results one may now prove

THEOREM 16.3. *A necessary and sufficient condition that*

$$f(x) = \int_0^\infty \frac{\phi(t)}{(x+t)^2}\,dt, \tag{16.6}$$

where $\phi(t)$ is bounded is that there should exist a constant N for which

$$|M_{k,t}[f(x)]| < N \qquad (k = 1, 2, 3, \cdots; 0 < x < \infty). \tag{16.7}$$

Note the contrast of this result with Theorem 15.1 by reason of the absence of any conditions of the type (15.3), (15.4). The proof follows by use of Theorems 16.1 and 16.2, and is omitted.

For the applications to follow it is desirable that the conditions of Theorem 16.3 involving the operator $M_{k,t}[f(x)]$ should be replaced by one involving $L_{k,t}[f(x)]$. We thus prove

THEOREM 16.4. *A necessary and sufficient condition that $f(x)$ should have the representation* (16.6) *with $\phi(t)$ bounded and satisfying*

$$\int_0^t \phi(u)du \sim At \qquad (t \to 0) \tag{16.8}$$

for some constant A is that

$$\left|\int_0^t L_{k,u}[f(x)]du\right| < N \qquad (k = 1, 2, 3, \cdots) \tag{16.9}$$

for some constant N.

If $f(x)$ has the representation described, then it follows from an abelian theorem, easily proved, that

$$f^{(k)}(x) \sim \frac{(-1)^k k! A}{x^{k+1}} \qquad (x \to 0). \tag{16.10}$$

But (16.10) shows that

$$\int_0^t L_{k,u}[f(x)]du = M_{k,t}[f(x)] - (-1)^{k-1}\left(\frac{k-1}{k}\right)A \tag{16.11}$$

$$(k = 2, 3, \cdots),$$

$$\int_0^t L_{1,u}[f(x)]du = M_{1,t}[f(x)] - A. \tag{16.12}$$

By Theorem 16.3 the right-hand sides of these equations have upper and lower bounds independent of k, so that the necessity of (16.9) is established.

Conversely, the existence of the integrals (16.9) implies by Theorem 10.2 the existence of a constant A such that

$$f^{(k)}(x) \sim \frac{(-1)^k A\, k!}{x^{k+1}} \qquad (x \to 0).$$

Hence (16.11) and (16.12) are again true and (16.9) implies (16.7). That is, (16.1) holds with $\phi(t)$ bounded. It remains only to establish (16.8). But this follows from (16.10), $k=0$, and the boundedness of $\phi(t)$ by a known Tauberian theorem.†

By use of these preliminary results we now prove

THEOREM 16.5. *A necessary and sufficient condition that*

$$f(x) = \int_0^\infty \frac{d\alpha(t)}{x+t},$$

where $\alpha(t)$ is a normalized function of bounded variation in every finite interval and is bounded in the infinite interval, is that there should exist a constant M and a positive function $N(t)$ such that

$$\left|\int_0^R L_{k,t}[f(x)]dt\right| \leqq M \qquad (R > 0;\ k = 1, 2, 3, \cdots), \tag{16.13}$$

$$\int_0^R |L_{k,t}[f(x)]|\, dt \leqq N(R) \qquad (R > 0;\ k = 1, 2, 3, \cdots). \tag{16.14}$$

If $f(x)$ has the representation described, then

$$f(x) = \int_0^\infty \frac{\alpha(t)}{(x+t)^2}\, dt, \tag{16.15}$$

$$\int_0^t \alpha(u)du \sim \alpha(0+)t \qquad (t \to 0),$$

† G. H. Hardy and J. E. Littlewood, *On Tauberian theorems*, Proceedings of the London Mathematical Society, vol. 30 (1930), p. 33, Theorem 5.

so that (16.3) follows from Theorem 16.4. Also

$$\lim_{k\to\infty}\int_0^R |L_{k,t}[f(x)]|\,dt = V(R) - V(0+)$$

from Theorem 6.3, so that the existence of $N(R)$ is insured.

Conversely, (16.13) implies that $f(x)$ has the form (16.15) with $\alpha(t)$ bounded and

$$\int_0^t \alpha(u)du \sim At \qquad (t\to 0).$$

Set

$$\alpha_k(t) = \int_0^t L_{k,u}[f(x)]du = M_{k,t}[f(x)] - (-1)^{k-1}\frac{k-1}{k}A.$$

But

$$M_{k,t}[f(x)] = d_k\int_0^\infty \frac{u^{k-1}t^k}{(t+u)^{2k}}\alpha(u)du,$$

and we showed in §3 that this integral approaches $\alpha(t)$ except perhaps in a set E of measure zero. But the variation of $\alpha_k(t)$ in $(0, R)$ is clearly not greater than $N(R)$ by (16.14). Hence $\alpha(t)$ is a normalized function of bounded variation in $(0, R)$ if suitably redefined. This redefinition has no effect on $f(x)$ since E is of measure zero. But

$$\int_0^\infty \frac{\alpha(t)}{(x+t)^2}dt = -\frac{\alpha(t)}{x+t}\Big|_0^\infty + \int_0^\infty \frac{d\alpha(t)}{x+t}.$$

The first term on the right-hand side is zero since $\alpha(t)$ is bounded and $\alpha(0)=0$. Hence the theorem is completely established.

It might be supposed that we could remove the restriction of boundedness of $\alpha(t)$ by considering the function

$$F(x) = \frac{f(x)-f(\delta)}{\delta - x}$$

as was done in §12. For even if $\alpha(t)$ becomes infinite, $F(x)$ satisfies (16.13) and (16.14) when the integral

$$f(x) = \int_0^\infty \frac{d\alpha(t)}{x+t} \tag{16.16}$$

converges. The converse is not true. If $F(x)$ satisfies (16.13) and (16.14) we have indeed

$$F(x) = \frac{f(x) - f(\delta)}{\delta - x} = \int_0^\infty \frac{\beta_\delta(t)}{(x+t)^2}\,dt,$$

where $\beta_\delta(t)$ is bounded. If $f(x)$ had the representation (16.16), then

$$\beta_\delta(t) = \int_0^t \frac{d\alpha(t)}{\delta + t}\cdot$$

But if $\alpha(t) = t \sin t$, then $\beta_\delta(t)$ is bounded. For

$$\beta_\delta(t) = \frac{t \sin t}{\delta + t} + \int_0^t \frac{u \sin u}{(\delta + u)^2}\,du,$$

and

$$\int_0^\infty \frac{u \sin u}{(\delta + u)^2}\,du$$

converges. But by Theorem 1.2 the integral (16.16) can not converge unless $\alpha(t) = o(t)$, which is not the case in the example considered.

17. **The Paley-Wiener inversion operator.** We conclude by showing the relation between the operator $L_{k,t}[f(x)]$ and the inversion operator given by Paley and Wiener.† They showed that if

$$f(x) = \int_0^\infty \frac{\phi(t)}{x+t}\,dt,$$

where $\phi(t)$ is a function of class L^2 in the interval $(0, \infty)$, then

$$\phi(t) = \operatorname*{l.i.m.}_{m\to\infty} \frac{1}{\pi t^{1/2}} \sum_{n=0}^{m} \frac{(-1)^n}{(2n)!}\left(\pi t \frac{d}{dt}\right)^{2n} (t^{1/2} f(t)).$$

We may abbreviate this precise result by the symbolic equation

$$\phi(t) = \frac{1}{\pi t^{1/2}} (\cos \pi \mathfrak{D})(t^{1/2} f(t)),$$

where

$$\mathfrak{D} = t\frac{d}{dt}\cdot$$

On the other hand

$$L_{k,t}[f(x)] = g_k \frac{1}{\pi t^{1/2}} \prod_{n=-k}^{k-2} \left(1 - \frac{2\mathfrak{D}}{2n+1}\right)(t^{1/2} f(t)), \tag{17.1}$$

where

† For reference see Introduction.

$$(17.2) \qquad g_k = \frac{\pi}{2}\left(\frac{1}{2}\cdot\frac{3}{2}\cdot\frac{3}{4}\cdot\frac{5}{4}\cdot\frac{5}{6}\cdot \cdots \cdot\frac{2k-5}{2k-4}\cdot\frac{2k-3}{2k-4}\right)\frac{2k-3}{2k-2}\cdot\frac{2k-1}{2k}.$$

To prove this one has only to verify that the differential operators on opposite sides of equation (17.1) have the same system of fundamental solutions,

$$\frac{1}{x^k}, \frac{1}{x^{k-1}}, \cdots, \frac{1}{x}, 1, x, \cdots, x^{k-3}, x^{k-2},$$

and to compare the coefficients of $f^{(2k-1)}(t)$ in the expanded forms of both operators. This coefficient for the left-hand side of (17.1) is

$$c_k(-1)^{k-1}t^{2k-1},$$

while for the right-hand side it is

$$\frac{g_k}{\pi}\prod_{n=-k}^{k-2}\left(\frac{-2}{2n+1}\right)t^{2k-1}.$$

Equating these coefficients gives (17.2).

Since

$$\cos \pi z = \lim_{k\to\infty}\prod_{n=-k}^{k-2}\left(1-\frac{2z}{2n+1}\right),$$

and

$$\frac{2}{\pi} = \prod_{n=1}^{\infty}\left(\frac{2n-1}{2n}\cdot\frac{2n+1}{2n}\right),$$

it follows that

$$\lim_{k\to\infty} g_k = 1,$$

and that

$$\lim_{k\to\infty} L_{k,t}[f(x)] = \frac{1}{\pi t^{1/2}}(\cos \pi D)(t^{1/2}f(t)),$$

so that the two operators are symbolically equivalent.

Harvard University,
Cambridge, Mass.

ANALYTICAL GROUPS*

BY
GARRETT BIRKHOFF

Introduction

1. **Abstract groups.** The present paper will deal with *abstract* continuous groups. This means that it will discuss symbols which behave like transformations, without specifying the domain on which the transformations operate. The reader will be assumed to be conversant with abstract groups as algebraic entities.

2. **Questions in the large.** It is well-known that the theory of continuous groups in the large differs essentially from the theory in the small. Some things, such as the one-one correspondence between closed subgroups of a Lie group and subalgebras of the Lie algebra of its infinitesimal generators, are true only locally;† others, such as the introduction by Weyl and Haar of invariant mass, are possible only when one deals with groups in the large.

The present paper is a theory *in the small* exclusively; it neither involves implicitly nor resolves explicitly the difficulties in the large. In this it resembles the original theory of Lie.

3. **Actual contents.** Thus the paper avoids two large classes of questions. What questions does it answer—what are its assumptions, and how can one summarize its conclusions?

The paper deals with systems (called "analytical groups") in which an associative multiplication is defined, and which can be so mapped on a Banach parameter-space that if one multiplies all elements by any fixed element near the origin, vector differences are left nearly invariant.‡

It is proved that if G is any analytical group (more properly, analytical group nucleus), then

(1) G is a topological group nucleus in the usual sense.

(2) One can introduce canonical parameters into G.

(3) G has an infinitesimal (Lie) algebra $L(G)$.

(4) The analytical subgroups of G correspond biuniquely to the closed

* Presented to the Society, December 31, 1936; received by the editors December 17, 1936 and, in revised form, May 3, 1937.

† Again, the group of topological automorphisms of the group of the torus differs radically from that of the group of translations of the plane, in spite of the fact that these two groups are locally isomorphic.

‡ A Banach space is of course simply a system having certain prescribed elementary properties of euclidean space which are shared by various important function spaces. Cf. §8.

subalgebras of $L(G)$, a subgroup being normal if and only if the corresponding subalgebra is invariant.

(5) If G is under canonical parameters, then there exists a formal series of polynomials determined by $L(G)$ which expresses the rule for forming group products.

(6) One can define product integrals for functions with values in $L(G)$, which include the Lebesgue integral (the case G is the additive group of real numbers), and all known product integrals (the case G is a group of matrices).

(7) Quite general functions $x(\lambda)$ with real arguments and values in a Banach space B determine formal series in elements of B and their brackets, which express the product integral $\hat{\int} x(\lambda)d\lambda$ under canonical parameters for *any* analytical group G whose parameter-space contains B and whose "commutation modulus" is sufficiently small.

(8) All the operations defined (e.g., vector addition under canonical parameters, product integration) are *topologico-algebraic*—preserved under topological isomorphisms.

4. **Extension to infinite dimensions.** Perhaps the main advantage of the above assumptions, is the fact that many *infinite* continuous groups satisfy them. This marks a real advance in the analytical theory of groups.*

The infinite-dimensional analytical groups treated in the literature are of two kinds: the infinite continuous groups of analytical transformations

$$x_i' = f_i(x_1, \cdots, x_n) \qquad (i = 1, \cdots, n)$$

of n-dimensional space discussed by Lie [10]† and Cartan [4], and the groups of linear operators on Hilbert space recently studied by Delsarte [6]. Each of these authors omits to define the meaning of the convergence $T_n \to T$ of a sequence of transformations T_n to a limit T—in other words, to define the topological structure of the corresponding abstract groups.

This omission, and the omission to establish a rigorous correlation between the actual transformations of such groups and so-called "infinitesimal" transformations, are not trivial. In fact, although the present paper supplies a complete theory for a class of groups including those studied by Delsarte, the author does not even know what the facts are in the case of the groups studied by Lie and Cartan. Part of the difficulty is that the group manifolds are not metric-linear; part of it is that canonical parameters do not define even locally a one-one representation of the group manifold.

* Cf. Abstract 41-3-129 of the Bulletin of the American Mathematical Society (1935); also *Continuous groups and linear spaces*, Matematicheskiĭ Sbornik, vol. 1 (1936), pp. 635–642; an address delivered at the First International Topological Conference, Moscow, September 5, 1935.

† Numbers in brackets refer to the bibliography at the end of the paper.

5. **Continuous groups: topological and analytical.** This illustrates the importance of geometrical properties of group manifolds; we shall now see how continuous groups can be classified on a purely geometrical basis.

A "continuous" abstract algebra (whether group, ring, or field) is a system whose elements are simultaneously the points of a geometrical manifold and the symbols of a formal calculus, and whose algebraic operations determine "smooth" functions of the manifold into itself. By letting the geometry of the manifold suggest the proper definition of smoothness, one is led to a purely geometrical classification of continuous abstract algebras.

Thus with groups whose manifolds are general topological spaces, one naturally regards "smoothness" of the group operations as meaning that group multiplication and passage to the inverse are continuous in the topology of the group manifold. Such groups are called *topological*.

Similarly, with groups whose manifolds are n-dimensional analytical varieties, it is natural to assume that the group operations are analytical in the coordinates; this leads to the usual concept of a *Lie* group.

Now it is a remarkable fact, that two analytical systems which are continuous images of each other, are in general analytical images of each other. This seems to hold even in pure geometry: thus dimensionality, originally known to be invariant only under analytical transformations, is now realized to be a topological invariant. We shall extend the domain of validity of this principle below, by proving that *continuous* isomorphisms between Lie groups are necessarily *analytical*.*

6. **Groups as topological algebras.** The result just stated, combined with (8), suggests that one can develop a theory for analytical groups in which group multiplication and passage to the limit are the only notions introduced as undefined primitives.†

Indeed, this program is technically feasible: it is shown below that one *can* give topologico-algebraic definitions of analytical groups. But as the argument is really *metric*, it would be misleading to make it pseudo-topological—even though it is less analytical and more topologico-algebraic‡ than any

* Discontinuous (and hence non-analytical) isomorphisms exist; there is one between the group of translations of the line and the group of translations of the plane. To see this, form in each an independent basis with respect to linear combination with rational coefficients. (However, van der Waerden, Mathematische Zeitschrift, vol. 36 (1933), pp. 780–787, has shown that isomorphisms between compact semi-simple Lie groups are always analytical.) Conceivably two Lie groups which are isomorphic and have homeomorphic manifolds are eo ipso analytically isomorphic.

† Especially since O. Schreier [15] has obtained so much information about group manifolds by such a theory.

‡ Thus pure group algebra—especially that of commutation—is shown to yield many results (especially in Chaps. IV–V) which could not be obtained by general analytical methods.

previous reasoning yielding the same results.

All this relates to the well-known problem of determining the weakest analytical assumptions demonstrably equivalent to an assumption of unrestricted analyticity. The weakest assumption in the literature* (cf. [11]) is that the function of group multiplication has continuous *second* derivatives. It is shown below that if one assumes continuous *first* derivatives, then one can deduce the whole theory† of abstract Lie groups.

Chapter I. Technical machinery

7. **A remark on notation.** It will shorten the argument in the sequel to use the following notational conventions: $M(\lambda)$ for any positive function of a real variable λ such that $\lim_{\lambda\to 0} M(\lambda)=0$; $O(\lambda)$ for any such function satisfying $O(\lambda)\leqq K\cdot|\lambda|$ for some $K<+\infty$; $o(\lambda)$ for any such function satisfying $o(\lambda)\leqq|\lambda|\cdot M(\lambda)$ for some $M(\lambda)$. (The relation of the last two definitions to Landau's well-known o-O notation‡ is obvious.)

Thus let $\phi(x_1,\cdots,x_r)$ and $\psi(x_1,\cdots,x_r)$ be any two real-valued functions of the same (not necessarily numerical!) variables $x_1,\cdots,x_r$. By the preceding definition,

$$\phi(x_1,\cdots,x_r)\leqq M(\psi(x_1,\cdots,x_r))$$

means that given $\eta>0$, $\delta>0$ exists so small that $\psi(x_1,\cdots,x_r)<\delta$ implies $\phi(x_1,\cdots,x_r)<\eta$. The inequalities $\phi(x_1,\cdots,x_r)\leqq O(\psi(x_1,\cdots,x_r))$ and $\phi(x_1,\cdots,x_r)\leqq o(\psi(x_1,\cdots,x_r))$ have similar meanings.

It is obvious that in terms of this notation, the following substitutions are legitimate:

$$(7\alpha)\qquad O(\lambda)\quad\text{for}\quad o(\lambda),\quad\text{and}\quad M(\lambda)\quad\text{for}\quad O(\lambda).$$

$$(7\beta)\qquad M(\lambda)\quad\text{for}\quad M(O(\lambda)),\quad\text{and}\quad O(\lambda)\quad\text{for}\quad O(O(\lambda)).$$

$$(7\gamma)\qquad M(\lambda+\mu)\quad\text{for}\quad M(\lambda)+M(\mu).$$

$$(7\delta)\qquad M(\lambda)\quad\text{for}\quad M(\lambda)/M[1-M(\lambda)].$$

Thus if $\phi(x_1,\cdots,x_r)\leqq M(\psi_1(x_1,\cdots,x_r))+M(\psi_2(x_1,\cdots,x_r))$, then by (7$\gamma$), $\phi(x_1,\cdots,x_r)\leqq M(\psi_1(x_1,\cdots,x_r)+\psi_2(x_1,\cdots,x_r))$.

It goes without saying that the M-functions, o-functions and O-functions appearing in the text vary from group to group, and from inequality to inequality—although since only a finite number of such functions are used in

* Except when dealing with compact (von Neumann) and abelian (Pontrjagin) groups, where one need only assume that one has a topological group locally homeomorphic with euclidean space.

† The author announced this result in Abstract 41-5-192 (1935) of the Bulletin of the American Mathematical Society.

‡ Cf. G. H. Hardy, *Pure Mathematics*, 5th edition, Cambridge University Press, 1928, p. 448.

dealing with any one group, there exists a single $M(\lambda)$, $o(\lambda)$, and $O(\lambda)$ which works in all inequalities for that group.

8. **Formal definition of "analytical group."** The properties of "analytical groups" which will be assumed were indicated in §2; they can be stated explicitly as

DEFINITION 1. By an *analytical group* will be meant any region about the origin of a Banach space, in which an associative multiplication is defined for elements near the origin Θ, satisfying

$$x \circ \Theta = \Theta \circ x = x \quad \text{for all} \quad x. \tag{1}$$

$$|(xa - xb) - (a - b)| \leqq M(|x| + |b| + |a|) \cdot |a - b|. \tag{2'}$$

$$|(ay - by) - (a - b)| \leqq M(|a| + |b| + |y|) \cdot |a - b|. \tag{2''}$$

In words, the origin is the group-identity e, and vector differences are nearly invariant under group translations $T_y^x : a \to xay$. (By xy or $x \circ y$ is meant the group product of x and y.)

(By a *Banach space* is meant a B-space in the sense of Banach [1]—that is, a linear space in which an absolute value $|x|$ is defined which (1) is positive for $x \neq \Theta$, (2) satisfies the triangle inequality $|x+y| \leqq |x| + |y|$, (3) is multiplied identically by $|\lambda|$ under any scalar expansion $x \to \lambda x$, and (4) makes the space complete*—such that if $\lim_{m,n\to\infty} |x_m - x_n| = 0$, then x exists such that $\lim_{n\to\infty} |x - x_n| = 0$.)

9. **A topological group nucleus.** Combining (2′)–(2″), we get immediately,

$$|(xay - xby) - (a - b)| = M(|x| + |a| + |b| + |y|) \cdot |a - b|. \tag{2}$$

Again, setting $b = e = \Theta$ in (2′) and $a = e = \Theta$ in (2″), one obtains,

$$|x \circ y - (x + y)| \leqq M(|x| + |y|) \cdot |y|, \tag{3'}$$

$$|x \circ y - (x + y)| \leqq M(|x| + |y|) \cdot |x|, \tag{3''}$$

which can be combined into the single inequality

$$|x \circ a \circ y - (x + a + y)| = M(|x| + |a| + |y|) \cdot (|x| + |y|). \tag{3}$$

In words, near the origin group translations T_y^x differ little from the corresponding linear translations $L_y^x : a \to a + x + y$.

(9α) *Multiplication is continuous near* Θ.

Proof. If $|x|$, $|y|$, $|a|$ and $|b|$ are small, then

$$\begin{aligned} |(x \circ y) - (a \circ b)| &= |\{(x \circ y) - (a \circ y)\} + \{(a \circ y) - (a \circ b)\}| \\ &= O(|x - a| + |y - b|). \end{aligned}$$

---

* Incidentally, $\{x_n\}$ is metrically "fundamental" if and only if it is "fundamental" in the topologico-algebraic sense (of van Dantzig) that $\lim_{m,n\to\infty} x_m^{-1} \circ x_n = 0$. Cf. (9δ).

(9β) *Every sufficiently small element* x *has a unique inverse* x^{-1} *satisfying* $x\circ x^{-1}=x^{-1}\circ x=\Theta$, *and* $|x^{-1}|\leqq 2\cdot|x|$.

Proof. Suppose $M(5|x|)<\frac{1}{2}$. Define $y_0=\Theta$ and by induction $y_{n+1}=y_n-(xy_n)$. Then

$$|xy_{n+1}| = |(xy_{n+1}-xy_n)-(y_{n+1}-y_n)|, \quad \text{by definition,}$$
$$\leqq M(|x|+|y_n|+|y_{n+1}|)\cdot|xy_n|, \quad \text{by (2}'\text{)},$$

since $|y_{n+1}-y_n|=|-xy_n|=|xy_n|$. It follows by induction that $|xy_{n+1}|\leqq\frac{1}{2}^{n+1}\cdot|x|$, and $|y_{n+2}|\leqq(2-\frac{1}{2}^{n+1})\cdot|x|$—whence $M(|x|+|y_{n+1}|+|y_{n+2}|)<\frac{1}{2}$. Hence $\lim_{m,n\to\infty}|y_m-y_n|=0$, and so by completeness a y exists satisfying $|y|\leqq 2\cdot|x|$ and $\lim_{n\to\infty}|y-y_n|=0$. Now by (9α), $xy=\Theta$, and so y is a right-inverse of x. Similarly x has a left-inverse z with $zx=\Theta$. Moreover $y=(zx)y=z(xy)=z$; hence $y=z\equiv x^{-1}$ is a full inverse of x; its uniqueness follows since $xx'=\Theta$ implies $x'=(x^{-1}x)x'=x^{-1}(xx')=x^{-1}$, while $x''x=\Theta$ implies $x''=x''(xx^{-1})=(x''x)x^{-1}=x^{-1}$.

(9γ) *Passage to the inverse is continuous near* Θ.

Proof. Let $(x+u)$ be given. Substitute x^{-1} for y_0 and $x+u$ for x in the proof of (9β). By (9α), $(x+u)y\leqq 2\cdot|u|$ in a small enough neighborhood; hence in the construction of $(x+u)^{-1}$ by successive approximation,

$$|(x+u)^{-1}-x^{-1}|\leqq|y-y_0|=4\cdot|u|.$$

We can summarize (9α)–(9γ) in

Theorem 1. *Every analytical group contains a topological group nucleus in the usual sense.**

A topological space in which an associative multiplication is defined satisfying (9α)–(9γ) everywhere is called a topological group (cf. [15]).

(9δ) $|x^{-1}|\leqq|x|+o(|x|)$; *in fact*, $|x^{-1}+x|=o(|x|)$.

Proof. By (3′), $|x+x^{-1}|\leqq M(|x|+|x^{-1}|)\cdot|x|$; but by (9β), $M(|x|+|x^{-1}|)=M(|x|)$. Hence $x^{-1}=-x+u$, where $|u|=M(|x|)\cdot|x|=o(|x|)$, proving the result.

Digression on axiomatics: Setting $y=\Theta$ resp. $x=\Theta$ in (3′) resp. (3″), we obtain (1). Further, near Θ, (3′), (3″) make $by=a$ imply that $|y|$ is nearly $|b-a|$. Hence if we are dealing with a topological group nucleus, then (2′) and (2″) hold. (Proof: By symmetry, it suffices to prove (2′). But by (3′), writing $b^{-1}a=y$, whence $a=by$,

* Cf. B. L. van der Waerden, *Vorlesungen über kontinuierlichen Gruppen*, Göttingen, 1932. For the analogous notion of a Lie group nucleus (alias "germ," cf. [11]).

$$|(a-b)-y| = |by-b-y| \leqq M(|b|+|y|)\cdot|y|$$

$$|(xa-xb)-y| = |xby-xb-y| = M(|x|+|b|+|y|)\cdot|y|.$$

And so by the triangle law, since by continuity $M(|y|)=M(|a^{-1}b|) \leqq M(|a|+|b|)$ (cf. §7),

$$(2') \qquad |(xa-xb)-(a-b)| \leqq M(|x|+|b|+|a|)\cdot|a-b|\ .)$$

10. **Groups in the large.** Let H be any full topological group. Obviously any one-one bicontinuous map of a neighborhood of the identity of H onto a region of a Banach space which satisfies (1), (2′), (2″)—or alternatively, by the last paragraph, (3′), (3″)—defines that neighborhood as an analytical group. We are unable to prove* that any system satisfying Definition 1 is conversely a piece containing the identity of a full topological group.

Full topological groups which can be mapped locally onto Banach space in such a way as to satisfy Definition 1 will be termed *full* analytical groups; this will distinguish full groups from the analytical group nuclei with which we shall be concerned below and which, for brevity, we have called simply "analytical groups."

11. **Changes of parameters.** It is important to know which transformations of Banach spaces play the role of analytical coordinate transformations in the theory of abstract Lie groups—that is, which when applied to a given analytical group G attached to a parameter space, turn G into another topologically isomorphic analytical group H.

One can specify at once two classes of such transformations associated with an arbitrary Banach space B, namely:

(11a) The group of "distortions" of B—that is, of those transformations

---

* This has been proved for finite continuous groups by E. Cartan ([5], p. 18). Cartan omits to mention the decisive fact that if L is any Lie algebra, and N is its largest invariant "integrable" subalgebra, then L contains a semi-simple subalgebra S such that $S \cap N=0$ and $S+N=L$ (cf. J. H. C. Whitehead, Proceedings of the Cambridge Philosophical Society, vol. 32 (1936), pp. 229–238). This omission led Mayer-Thomas to question ([11], p. 806) the validity of Cartan's proof. Cartan has since published another equally technical proof (*Sur la Topologie des Groupes de Lie*, Paris, 1936, p. 22).

Neither of these proofs can be extended to the infinite continuous group nuclei treated below; each depends on lemmas which need not be true in infinite dimensions. For instance the fact that not all closed linear subspaces S of Banach spaces B have complements T such that $S \cap T=0$ and $S+T=B$ prevents one from using Cartan's special proof for solvable groups.

On the other hand Mayer-Thomas' argument (due independently to Paul Smith) for the case of group nuclei which can be embedded in a full group generalizes to infinite continuous groups—one takes the subgroup of the full group generated by the nucleus given, and retopologizes this subgroup by redefining distance as geodesic distance along paths in the subgroup.

Esthetically, one would expect to find a simple proof that every analytical group nucleus can be embedded in a full group, since it is easy to define the full group, if one knows that it exists.

T of B into itself which leave Θ fixed and satisfy (*) $|T(a+x)-T(a)-x| \leqq M(|a|+|x|)\cdot|x|$.

(11b) The class of alterations of the norm function $|x|$ of B to a new norm function $\|x\|$ such that the ratios $|x|/\|x\|$ and $\|x\|/|x|$ are uniformly bounded.

Remark. The latter correspond one-one to choices of bounded open convex regions $\|x\|<1$ of B. (Cf. A. Kolmogoroff, op. cit. in §17.)

THEOREM 2. *Any succession of transformations of types* (11a), (11b) *of the parameter-space of an analytical group G, turns G into a topologically isomorphic analytical group.*

Proof. It is obviously sufficient to prove the theorem for single transformations of types (11a) and (11b); again, the main difficulty is to prove analyticity. With (11b), one needs merely write $|x|/R \leqq \|x\| \leqq R\cdot|x|$, and to replace $M(\lambda)$ in (3) by $R^2M(\lambda/R) \leqq M(\lambda)$. (Cf. §7.)

Consider case (11a). Setting $a+x=b$ in (*), we see that (**) $|T(b)-T(a)-(b-a)| \leqq M(|a|+|b|)\cdot|b-a|$—i.e., vector differences, and hence absolute values near Θ are nearly invariant under T. Hence—the proof as in (9β) is by successive approximation—T is one-one and so by (**) bicontinuous near the origin. Therefore it suffices to prove (3′), (3″)—or even, by symmetry, (3′). This we shall do. Note that T^{-1} is of type (11a), and leaves vector differences near Θ almost invariant. Hence

$$|T^{-1}(a+x)-(T^{-1}(a)+T^{-1}(x))| \leqq M(|a|+|x|)\cdot|x|, \text{ by } (*),$$

$$|T^{-1}(a)\circ T^{-1}(x)-(T^{-1}(a)+T^{-1}(x))| \leqq M(|a|+|x|)\cdot|x|, \text{ by } (3').$$

Hence by the triangle inequality,

$$|T^{-1}(a)\circ T^{-1}(x)-T^{-1}(a+x)| \leqq M(|a|+|x|)\cdot|x|$$

and so, by (**),

$$|T(T^{-1}(a)\circ T^{-1}(x))-(a+x)| \leqq M(|a|+|x|)\cdot|x|.$$

But this *is* (3′) in terms of the new parameters, q.e.d.

We shall regard topologically isomorphic groups as essentially identical—as differing merely in their parametric representation.

12. **Rectifiable paths.** Let us recall a few familiar geometrical notions, so as to have a consistent notation and terminology for subsequent use. These notions are proper to abstract metric spaces,† and so apply to Banach spaces.

By a *path* is meant a continuous image P: $p(\lambda)$ of a line interval‡ $[0, \Lambda]$.

† The ideas go back to Fréchet's Thesis; cf., also, K. Menger, *Zur Metrik der Kurven*, Mathematische Annalen, vol. 103 (1930), p. 471, §5.

‡ As is conventional, $[0, \Lambda]$ denotes the set of real numbers λ which satisfy $0 \leqq \lambda \leqq \Lambda$

Two paths P and Q are called "geometrically equivalent" (written $P\approx Q$) if and only if they can be identified by proper choice of parameters—i.e., if and only if one can establish a one-one sense-preserving correspondence between the intervals of which they are images, such that corresponding points have the same image. Clearly the relation of being geometrically equivalent is reflexive, symmetric and transitive.

Again, by a *segment* ΔP of P is meant the image of any subinterval Δ: $[\lambda_1, \lambda_2]$ of $[0, \Lambda]$. By a *partition* π of P is meant a division of $[0, \Lambda]$ into successive subintervals Δ_k: $[\lambda_{k-1}, \lambda_k]$, where $\lambda_0=0$, $\lambda_n=\Lambda$, and $k=1, \cdots, n$. By the "product" of any two partitions π and π' of P is meant the partition $\pi\cdot\pi'$ whose subintervals are the intersections of the subintervals of π with those of π'. And π is called a "subpartition" of π' (in symbols, $\pi\leqq\pi'$) if and only if $\pi=\pi\cdot\pi'$.

By the π-approximate length of P under any partition π is meant $|P|_\pi\equiv\sum_{k=1}^{n}|P(\lambda_k)-P(\lambda_{k-1})|$, and by the "length" of P is meant $|P| =\sup |P|_\pi$. A path P is called "rectifiable" if and only if $|P| <+\infty$. Obviously

(12α) *If* $P\approx Q$, *then* $|P| = |Q|$.

The "diameter" $|\pi|$ of a partition π of a rectifiable path P is defined as $\sup |\Delta P|$. It is not hard to show

(12β) $|P| =\lim_{|\pi|\to 0}|P|_\pi$.

And since $|\pi| \leqq |\pi'|$ provided $\pi\leqq\pi'$, we see

(12γ) $|P| =\lim_{\pi\downarrow}|P|_\pi$ *in the sense of Moore-Smith.*†

13. **More notation.** We shall now introduce some special but natural notation for handling *rectifiable paths issuing from the origin* (=identity) *of an analytical group nucleus.*

If P_k is any path with domain $[0, \Lambda_k]$, then $t(P_k)$ denotes $p_k(\Lambda_k)-p_k(0)$.

By the *path-sum* of r admissible paths $P_1, \cdots, P_r$, will be meant the path $P=P_1\oplus \cdots \oplus P_r$ formed by adding to $P_1\oplus \cdots \oplus P_{r-1}$ a segment congruent to P_r under linear translation through $t(P_1\oplus \cdots \oplus P_{r-1})$. And by the *path-product* of the P_k, will be meant the path $\tilde{P}=P_1\circ \cdots \circ P_r$ formed by adding to $P_1\circ \cdots \circ P_{r-1}$ a segment congruent to P_r under group left-translation through $t(P_1\circ \cdots \circ P_{r-1})$. Thus P and $\tilde{P}$ have $[0, \Lambda_1+ \cdots +\Lambda_r]$ for domain, and for $0\leqq\lambda\leqq\Lambda_{k+1}$,

$$(13.1)\qquad \begin{cases} p(\Lambda_1 + \cdots + \Lambda_k + \lambda) = t(P_1) + \cdots + t(P_k) + p_{k+1}(\lambda), \\ \tilde{p}(\Lambda_1 + \cdots + \Lambda_k + \lambda) = t(P_1) \circ \cdots \circ t(P_k) \circ p_{k+1}(\lambda). \end{cases}$$

† This means that, given any neighborhood of $|P|$, one can find a π_0 such that $\pi\leqq\pi_0$ implies that $|P|_\pi$ lies in that neighborhood of $|P|$.

Since linear and group translations leave distances invariant resp. nearly invariant, P and $\tilde{P}$ are admissible.

We shall now develop an abstract correspondence between paths which generalizes the correspondence between the sum-integral $\int x(\lambda)d\lambda$ and the product integral $\widehat{\int} x(\lambda)d\lambda$ of functions whose values $x(\lambda)$ are matrices. (N.B., $t(\Delta_k P)$ is the analogue of $x(\lambda)\Delta\lambda$.)

Accordingly, let P be any admissible path, and let π be any partition of P into segments $\Delta_1 P, \cdots, \Delta_r P$. Denote by P_k the image of $\Delta_k P$ after linear translation through $-t(P_1 \oplus \cdots \oplus P_{k-1})$, and by Q_k the image of $\Delta_k P$ after left-multiplication by the group-inverse of $t(Q_1 \circ \cdots \circ Q_{k-1})$. Then by construction

$$P = P_1 \oplus \cdots \oplus P_r = Q_1 \circ \cdots \circ Q_r.$$

We shall define the two dualistic paths

$$P_\pi^* = P_1 \circ \cdots \circ P_r \qquad \text{and} \qquad P_\pi \dagger = Q_1 \oplus \cdots \oplus Q_r$$

formed by interchanging the operations of path-addition and path-multiplication. Then we shall prove that the P_π^* and $P_\pi\dagger$ approach fixed limiting positions P^* and $P\dagger$ as $|\pi|$ tends to zero.

14. **Evaluation of paths.** Of course, the meaning of this statement depends on how one defines limits—on how one *topologizes* the "space" of images of a fixed interval.

Let P and Q be any two images of the same interval $[0, \Lambda]$. We shall make the definition

$$|P - Q| = \sup_{0 \leq \lambda \leq \Lambda} p(\lambda) - q(\lambda).$$

It is clear that this definition of distance makes the images of $[0, \Lambda]$ the "points" of an abstract metric space, in the sense defined earlier;‡ this depends only on the fact that the images of $[0, \Lambda]$ are themselves in a metric space.

We now come to some statements involving group properties. In stating and proving them we shall write $\prod_{k=1}^r x_k$ for $(x_1 \circ \cdots \circ x_r)$, and $\sum_{k=1}^r x_k$ for $(x_1 + \cdots + x_r)$.

(14α) $\left|\prod_{k=1}^r x_k - \sum_{k=1}^r x_k\right| = o(\sum_{k=1}^r |x_k|)$. *Consequently,* $\left|\prod_{k=1}^r x_k\right| \leq O(\sum_{k=1}^r |x_k|)$.

Proof. By the triangle inequality,

$$\left|\prod_{k=1}^r x_k - \sum_{k=1}^r x_k\right| \leq \left|\left[\left(\prod_{k=1}^{r-1} x_k\right) \circ x_r\right] - \left[\left(\prod_{k=1}^{r-1} x_k\right) + x_r\right]\right|$$

‡ Fréchet, op. cit., p. 36, introduces this very definition of distance, and shows that it is metric.

$$+\left|\prod_{k=1}^{r-1} x_k - \sum_{k=1}^{r-1} x_k\right|$$

$$\leqq |x_r| \cdot M\left(\left|\prod_{k=1}^{r-1} x_k\right|\right) + M\left(\sum_{k=1}^{r-1}|x_k|\right)\cdot\left(\sum_{k=1}^{r-1}|x_k|\right)$$

by (3) and induction on r. Recombining—since, by induction on r, $|\prod_{k=1}^{r-1}x_k| \leqq O(\sum_{k=1}^{r-1}|x_k|)$—we get

$$\left|\prod_{k=1}^{r} x_k - \sum_{k=1}^{r} x_k\right| \leqq M\left(\sum_{k=1}^{r-1}|x_k|\right)\left(\sum_{k=1}^{r-1}|x_k|\right) \leqq o\left(\sum_{k=1}^{r}|x_k|\right).$$

(14β) *We have the inequality*

$$\left|\prod_{k=1}^{r} x_k - \prod_{k=1}^{r} y_k - \sum_{k=1}^{r}(x_k - y_k)\right| \leqq M\left(\sum_{k=1}^{r}|x_k| + \sum_{k=1}^{r}|y_k|\right)$$

$$\cdot\left(\sum_{k=1}^{r}|x_k - y_k|\right).$$

Hence if $M(\sum_{k=1}^{r}[|x_k|+|y_k|])<1$, *then*

$$\left|\prod_{k=1}^{r} x_k - \prod_{k=1}^{r} y_k\right| \leqq 2\left(\sum_{k=1}^{r}|x_k - y_k|\right).$$

Proof. By the triangle inequality iterated,

$$\left|\prod_{k=1}^{r} x_k - \prod_{k=1}^{r} y_k - \sum_{k=1}^{r}(x_k - y_k)\right| \leqq \sum_{k=1}^{r}\left|\left(\prod_{i=1}^{k-1} x_i\right)x_k\left(\prod_{i=k+1}^{r} y_i\right)\right.$$

$$\left. - \left(\prod_{i=1}^{k-1} x_i\right)y_k\left(\prod_{i=k+1}^{r} y_i\right) - (x_k - y_k)\right|$$

$$\leqq \sum_{k=1}^{r} M\left(\sum_{k=1}^{r}|x_k| + \sum_{k=1}^{r}|y_k|\right)\cdot|x_k - y_k|, \qquad \text{by (2)},$$

since $|\prod_{i=1}^{k-1}x_i| \leqq O(\sum_{i=1}^{r}|x_i|)$ and $|\prod_{i=k+1}y_i| \leqq O(\sum_{i=1}^{r}|y_i|)$ by (14α).

(14γ) *Let* $P_1, \cdots, P_r$ *and* $Q_1, \cdots, Q_r$ *be admissible paths, each* P_k *having the same domain as* Q_k. *Further, let* $|P|$ *denote* $\sum_{k=1}^{r}|P_k|$ *and* $|Q|$ *denote* $\sum_{k=1}^{r}|Q_k|$. *Then*

$$|(P_1 \oplus \cdots \oplus P_r) - (Q_1 \oplus \cdots \oplus Q_r)| \leqq \sum_{k=1}^{r}|P_k - Q_k|.$$

And if $|P|+|Q|$ *is so small that* $M(|P|+|Q|)<1$, *then*

$$|(P_1 \circ \cdots \circ P_r) - (Q_1 \circ \cdots \circ Q_r)| \leqq 2\sum_{k=1}^{r}|P_k - Q_k|.$$

Proof. The first inequality follows from (13.1) and the triangle inequality. The second follows from (13.1) and (14β).

Thus with paths of sufficiently small total length, both path-sums and path-products are uniformly continuous functions of their arguments in our metric "path-space."

LEMMA. *Let P be any sufficiently short path. Then if $\pi' \geqq \pi$, $|P_\pi^* - P_{\pi'}^*| \leqq M(|\pi|)$ and $|P_\pi\dagger - P_{\pi'}\dagger| \leqq M(|\pi|)$.*

Proof. It is an essential preliminary remark that each segment of any P_π^* or $P_\pi\dagger$ has nearly the same length as the corresponding segment of P, since it is obtained from it by group and linear translation of subsegments through relatively small distances—and such translations by (2) leave distances nearly invariant.

Now write $P_\pi^* = P_1 \circ \cdots \circ P_r$. Clearly $P_{\pi'}^*$ is obtainable from P_π^* by replacing each component $P_k = P_{k,1} \oplus \cdots \oplus P_{k,s}$ by the path $\tilde{P}_k = P_{k,1} \circ \cdots \circ P_{k,s}$. But referring to (14$\alpha$), we see that $|\tilde{P}_k - P_k| \leqq o(|P_k|) \leqq M(|\pi|) \cdot |P_k|$. Hence by (14$\gamma$),

$$| P_\pi^* - P_{\pi'}^* \leqq 2M(|\pi|) \cdot \sum_{k=1}^{r} |P_k| = M(|\pi|) \cdot |P|.$$

Similarly, write $P_\pi\dagger = Q_1 \oplus \cdots \oplus Q_r$. Clearly $P_{\pi'}\dagger$ is obtainable from $P_\pi\dagger$ by replacing each component $\tilde{Q}_k = Q_{k,1} \circ \cdots \circ Q_{k,s}$ by a path $\tilde{Q}_k = Q_{k,1} \oplus \cdots \oplus Q_{k,s}$. By (14$\alpha$) and the preliminary remark, $|\tilde{Q}_k - Q_k| \leqq o(|Q_k|) \leqq M(|\pi|) \cdot |Q_k|$. Hence by (14$\gamma$),

$$| P_\pi\dagger - P_{\pi'}\dagger | \leqq M(|\pi|) \cdot \sum_{k=1}^{r} |Q_k| = M(|\pi|) \cdot |P|.$$

THEOREM 3. *Let P be any sufficiently short path. Then paths P^* and $P\dagger$ exist such that*

$$| P_\pi^* - P^* | \leqq M(|\pi|) \qquad \text{and} \qquad | P_\pi\dagger - P\dagger | \leqq M(|\pi|).$$

Proof. By the above lemma, the P_π^* and $P_\pi\dagger$ converge in the sense of Cauchy-Fréchet. But this means that for every fixed λ, the $p_\pi^*(\lambda)$ and $p_\pi\dagger(\lambda)$ do, and hence (the space being complete) have limits $p^*(\lambda)$ and $p\dagger(\lambda)$. These limits define P^* and $P\dagger$; the inequalities of Theorem 3 then follow from the corresponding inequalities in the lemma and passage to the limit.

(14δ) $(P^*)\dagger = (P\dagger)^* = P$.

Proof. For every partition π, $(P_\pi^*)_\pi\dagger = (P_\pi\dagger)_\pi^* = P$ by definition. And to replace each segment of P_π^* or $P_\pi\dagger$ by the corresponding segment of P^* resp. $P\dagger$ makes by (14γ) a proportionally small change in $(P_\pi^*)_\pi\dagger$ resp. $(P_\pi\dagger)_\pi^*$. Hence $(P^*)_\pi\dagger \to P$ and $(P\dagger)_\pi^* \to P$ uniformly as $|\pi| \to 0$.

(14ε) $|t(P^*)-t(P)| \leqq o(|P|)$ *and* $|t(P\dagger)-t(P)| \leqq o(|P|)$.

Proof. For every π, $t(P_\pi^*)=t(P_1) \circ \cdots \circ t(P_r)$ where $t(P)=t(P_1)\oplus \cdots \oplus t(P)$. Hence the first inequality merely restates (14α). The proof of the second inequality is the same, since (cf. the preliminary remark in the proof of the lemma above) $|P\dagger| \leqq O(|P|)$.

$$(14\zeta) \quad (P_1\oplus \cdots \oplus P_r)^* = P_1^* \circ \cdots \circ P_r^*$$

and

$$(P_1 \circ \cdots \circ P_r)\dagger = P_1\dagger \oplus \cdots \oplus P_r\dagger.$$

Proof. $(P_1\oplus \cdots \oplus P_r)^*$ is in particular the limit as $\sup|P_{k,i}|\to 0$ of $P_{1,1} \circ \cdots \circ P_{r,s(r)}$, where

$$P_{k,1} \oplus \cdots \oplus P_{k,s(k)} = P_k.$$

Thus it is the limit as $\sup|\pi_k|\to 0$ of $(P_1)_{\pi_1}^* \circ \cdots \circ (P_r)_{\pi_r}^*$. By (14γ), this limit is $P_1^* \circ \cdots \circ P_r^*$. This proves the first identity; the proof of the second is similar.

Conversely (14ε)–(14ζ) define the correspondences $P\to P^*$ and $P\to P\dagger$.

(14η) *If Q is any path, and Δ: $[\lambda, \mu]$ is any interval of its domain, then* $q^{-1}(\lambda) \circ q(\mu) = t((\Delta Q\dagger)^*)$.

Remark. By $q^{-1}(\lambda)$ is of course meant $[q(\lambda)]^{-1}$.

Proof. Set $Q\dagger = P$; $p^*(\mu) = p^*(\lambda) \circ t((\Delta P)^*)$ by (14ζ); the result follows by transposing $p^*(\lambda) = q(\lambda)$.

(14θ) *If $P\approx Q$, then $P^*\approx Q^*$ and $P\dagger\approx Q\dagger$.*

Proof. Obvious from the definitions.

Chapter II. Canonical parameters

15. **Scalar powers.** In §§16–17, we shall consider straight rays P_x: $p_x(\lambda)=\lambda x$ $(0\leqq\lambda\leqq 1)$ and their star correspondents P_x^*.

Obviously $|P_x| = |x|$, and so by (14ε),

(15α) $|t(P_x^*)-x| \leqq o(|x|)$.

(15β) $|t(P_{x+y}^*)-t(P_x^*)-t(P_y^*)| \leqq M(|x|+|y|)\cdot|y|$.

Proof. Let π denote the partition of $[0, 1]$ into n equal parts. Then setting $x_k = x/n$ and $y_k=(x+y)/n$ in (14β), we get (15β) for $(P_{x+y})_\pi^*$, $(P_x)_\pi^*$ and $(P_y)_\pi^*$. Passing to the limit as $n\to\infty$, we get (15β).

Combining (15β) with (15α), we see that the so-called "canonical transformation" $T: x\leftarrow t(P_x^*)$ satisfies‡ $|T(x+y)-T(x)-y| \leqq M(|x|+|y|)\cdot|y|$—is of type (11a). Hence (cf. Theorem 2) we have

‡ By $T: x\leftarrow t(P_x^*)$ we mean that the position x is imagined to be occupied by the element $t(P_x^*)$.

THEOREM 4. *The canonical transformation carries any analytical group G into a topologically isomorphic analytical group.*

Again, by definition, $P_{(\lambda\mu)x}=P_{\lambda(\mu x)}$. While unless $\lambda\mu<0$, $P_{(\lambda+\mu)x}\approx P_{\lambda x}\oplus P_{\mu x}$, whence $t(P_{(\lambda+\mu)x})=t(P_{\lambda x}^*)\circ t(P_{\mu x}^*)$. But

(15γ) *Let* $R_x=P_x\oplus P_{-x}$. *Then* $t(R_x^*)=0$.

Proof. Let π denote the partition of R_x into $2n$ equal segments, and set $\lambda=1/n$. Then

$$\begin{aligned} |t((R_x)_\pi^*)| &= |(\lambda x)^{n-1}\circ(\lambda x\circ-\lambda x)\circ(-\lambda x)^{n-1}| \\ &\leqq |(\lambda x)^{n-1}\circ(-\lambda x)^{n-1}| + 2(|\lambda x\circ-\lambda x|) \end{aligned}$$

by (2), substituting $(\lambda x)^{n-1}$ for x, $(-\lambda x)^{n-1}$ for y, $(\lambda x\circ-\lambda x)$ for a and Θ for b, and requiring x to be so small that $M(|x|+|a|+|y|)<1$, whence

$$|xay| \leqq |xy| + |a| + M(|x|+|a|+|y|)\cdot|a| \leqq |xy| + 2|a|.$$

But by induction on n and (14α), this yields

$$\begin{aligned} |t((R_x)_\pi^*)| &= (n-1)\cdot o(|\lambda x|) + o(|\lambda x|) = n\cdot o(|\lambda x|) \\ &= n\cdot M(|\lambda x|)\cdot|\lambda|\cdot|x| = M(|\lambda x|)\cdot|x|. \end{aligned}$$

To complete the proof, let $n\to\infty$, so that $M(|\lambda x|)\to 0$.

But if $\lambda\mu<0$, $P_{\lambda x}\oplus P_{\mu x}\approx P_{(\lambda+\mu)x}\oplus P_{-\mu x}\oplus P_{\mu x}$; hence in all cases $t(P_{(\lambda+\mu)x})=t(P_{\lambda x}^*)\circ t(P_{\mu x}^*)$, and so we have

THEOREM 5. *For fixed x, the $t(P_{\lambda x}^*)$ are (locally) topologically isomorphic with the additive group of the λ.*

But by Theorem 4 the canonical transformation is one-one; hence the function x^λ defined by making $t(P_y^*)=x$ and $x^\lambda=t(P_{\lambda y}^*)$ is defined and single-valued near the origin.

(15δ) $x^1=x$ (by definition), $x^\lambda\circ x^\mu=x^{\lambda+\mu}$, *and* (since $P_{(\lambda\mu)x}=P_{\lambda(\mu x)}$) $(x^\lambda)^\mu=x^{\lambda\mu}$.

(15ε) x^λ *is a topologico-algebraic function of x, in the sense that any topological isomorphism carrying x into y carries x^λ into y^λ.*

Proof. The assertion is true for positive integral $\lambda=n$ since $(x\circ\cdots\circ x)=x^{1+\cdots+1}=x^n$. It is also true for positive rational λ since $y^n=x$ if and only if $y=(y^n)^{1/n}=x^{1/n}$; while since $x\circ y=\Theta$ if (by 15γ)) and only if (by (9β)) $y=x^{-1}$, it is true for all rational $\lambda=m/n$. Finally, since the rationals are dense in the real continuum and x^λ depends continuously on λ, it is true for all λ.

16. **Canonical parameters.** We are now in a position to introduce canonical parameters.

A group will be said to be under "canonical parameters" if and only if

the canonical transformation T: $x \leftarrow t(P_x^*)$ is the identical transformation I: $x \rightarrow x$.

(16α) *Any analytical group is transformed into canonical parameters by the canonical transformation—that is, the canonical transformation is idempotent.*

Proof. After T has been performed once, if π denotes the partition of (0, 1) into n equal parts, then by definition and Theorem 5, $t((P_x)_\pi^*) = (x/n)^n = x$, whence, passing to the limit, iteration of T leaves all points fixed.

(16β) *Under canonical parameters,* $x^\lambda = \lambda x$; *hence scalar multiplication under canonical parameters is an intrinsic topologico-algebraic operation.*

Proof. By definition of x^λ resp. canonical parameters, $x^\lambda = t(P_{\lambda x}^*) = \lambda x$.

(16γ) *In any analytical group,* $x+y = \lim_{\lambda\to 0}(\lambda x \circ \lambda y)/\lambda$.

Proof. Referring to the inequality (3), we get for fixed x and y since $\lambda(x+y) = \lambda x + \lambda y$,

$$|(\lambda x \circ \lambda y) - \lambda(x+y)| \leqq M(|\lambda|)\cdot|\lambda|.$$

Hence, dividing through by the scalar λ,

$$|(\lambda x \circ \lambda y)/\lambda - (x+y)| \leqq M(|\lambda|)$$

which completes the proof.

Combining (16γ) with (16β), we get

THEOREM 6. *If G and H are any two analytical groups under canonical parameters, then any topological isomorphism between G and H is linear—it preserves vector sums and scalar products.*

COROLLARY 6.1. *The group of topological automorphisms of any analytical group is spatially isomorphic with a group of linear transformations of its parameter-space.*

COROLLARY 6.2. *If G and H are any two analytical groups, then any continuous isomorphism between G and H can be expressed as the product of three transformations of the parameter-space of G, of types* (11a), (11b), *and* (11a).‡

COROLLARY 6.3. *Admissible paths* (cf. §13) *are carried into admissible paths under topological isomorphisms between analytical groups.*

(16δ) *An analytical group is under canonical parameters if and only if* $x \circ x = x + x$ *for all* x.

‡ One should prove further: *Any topological isomorphism between two groups whose function of composition is analytical, amounts to an analytical transformation of coordinates.* To complete the proof, it would suffice to show that in such groups $t(P_x^*)$ is an analytical function of x—a fact already known (from the theory of differential equations) for Lie groups.

Proof. If $x \circ x = x+x$, then by induction $x^{2^n} = 2^n x$, $x = (2^{-n}x)^{2^n}$, whence $t(P_x^*) = x$, and we have canonical parameters. Conversely, under canonical parameters, $x \circ x = x^2 = 2x = x+x$.

17. **Digression: Topologico-algebraic postulates.** It is a curious fact that, by inverting the remarks of the last few sections, one can obtain *topologico-algebraic postulates defining Lie groups*, involving only *intrinsic* operations (i.e., operations invariant under topological isomorphisms). To show this, one need use only superficial reasoning, arguing from the above properties of canonical parameters.†

One can do this even for infinite continuous groups. The general procedure is: 1°: characterize Banach spaces topologico-algebraically (as those complete topological linear spaces possessing a convex open "bounded" set‡); 2°: define linear transformations and thence Fréchet total derivatives (cf. §18) topologico-algebraically;§ 3°: postulate that the group is a Banach space relative to addition under canonical parameters ("canonical addition") and raising to scalar powers; 4°: postulate that an associative operation of multiplication satisfying $x \circ x = x+x$ and continuously differentiable on the Banach space, be defined.

Because of the preceding results and Corollary 2 of Theorem 15, these postulates are satisfied by all analytical groups under canonical parameters. Conversely, by Theorem 8 any system satisfying these postulates is an analytical group, which is by (16δ) under canonical parameters.

In the special case of Lie groups—the case that the parameter space has a finite basis (or equivalently,‖ is locally compact)—one can simplify these postulates to the requirements (i) elements $a_1, \cdots, a_r$ exist such that any element near the identity can be represented uniquely as a product $a_1^{\lambda_1} \circ \cdots \circ a_r^{\lambda_r}$ of small powers $a_k^{\lambda_k}$ of the a_k, and (ii) the function of composition is continuously differentiable¶ in $(\lambda_1, \cdots, \lambda_r)$-space.

18. **Digression: metric postulates.** The present section will be devoted to sketching a proof of

---

† These ideas were announced in Abstract 41-5-192 of the Bulletin of the American Mathematical Society (1935).

‡ For the terminology cf. J. von Neumann, *On complete topological spaces*, these Transactions, vol. 37 (1935), pp. 1–20. For the characterization cf. A. Kolmogoroff, *Zur Normierbarkeit eines allgemeinen topologisches lineares Raumes*, Studia Mathematica, vol. 5 (1935), pp. 29–33.

§ Replace the usual epsilon-delta definitions by "for every given neighborhood there exists a neighborhood so small $\cdots$."

‖ Cf. [1], p. 84, Theorem 8.

¶ This can be phrased topologico-algebraically. For instance, $x \circ y = f(x, y)$ has continuous first derivatives if and only if $\partial f/\partial x(a, b) = \lim_{\lambda \to 0}((a+\lambda x) \circ b)/\lambda$ and $\partial f/\partial y(a, b) = \lim_{\lambda \to 0}(a \circ (b+\lambda y))/\lambda$ exist and are continuous functions of a and b.

THEOREM 7. *One can redefine the class of analytical groups under canonical parameters by weakening the postulates for Banach spaces.*

This result will not be used elsewhere.

Sketch of proof. Substitute "group products" $x \circ y$ for vector sums $x+y$, and "scalar powers" x^λ for scalar products λx, continue to use an (extrinsic) norm function $|x|$, and make the following alterations in the usual postulates (cf. [1]) for Banach space (after first confining their validity to a small region about the identity): (1) replace the two conditions $x+y=y+x$ and $\lambda(x+y)=\lambda x+\lambda y$ by the single weaker condition $|x^\lambda \circ y^\lambda \circ (x \circ y)^{-\lambda}| \leqq |\lambda| \cdot |x| \cdot |y|$, and (2) replace the condition $|x+y| \leqq |x|+|y|$ by the weaker condition $|x \circ y| \leqq |x|+|y|+|x| \cdot |y|$.

The reader should have no difficulty in proving that the altered postulates hold in any analytical group under canonical parameters and under a suitable norm function (cf. Theorems 1 and 5 for the algebraic identities, and (27β)—where it is shown that essentially $|(x \circ y)-(x+y)| \leqq |x| \cdot |y|$—for the strong metric inequalities).

But conversely, if one defines $x+y=\lim_{\lambda\to 0}(x^\lambda \circ y^\lambda)^{1/\lambda}$ and $\lambda x=x^\lambda$ in any system G satisfying the new postulates, then the space becomes a neighborhood of the origin of a Banach space B, and the map of G on B satisfies (1), (2′), (2″) and $x \circ x=x+x$—completing the outline of the proof.

19. **Digression: differentiability postulates.** We now come to the connection between Definition 1 and differentiability conditions, namely

THEOREM 8. *Let G be any topological group nucleus, some neighborhood of whose identity e is mapped onto a region of a Banach space B, in such a way that $x \circ y=f(x, y)$ has first total derivatives everywhere, which are continuous at e. Then G is an analytical group under the map.*

Proof. Theorem 8 is clearly meaningless until continuous total derivatives have been defined; actually, it refers to the usual definitions due to Fréchet.† Fréchet says that $f(x, y)$ has a total derivative A with respect to x at $x=a$, $y=b$ if and only if there exists a linear transformation A such that

$$(18') \qquad |f(a+x, b)-f(a, b)-Ax| \leqq o(|x|),$$

where Ax denotes the transform of x by A. One similarly defines total derivatives with respect to y. Further, Fréchet calls the two total derivatives $A(x, y)=\partial f/\partial x(x, y)$ and $B(x, y)=\partial f/\partial y(x, y)$ *continuous at* $x=a$, $y=b$ if and only if

† M. Fréchet, *La notion de différentielles dans l'analyse générale*, Annales de l'École Normale Supérieure, (3), vol. 42 (1925), pp. 293–323. For a similar concept of an infinite continuous group, cf. A. D. Michal and V. Elconin, *Abstract transformation groups*, American Journal of Mathematics, vol. 59 (1937), pp. 129–144.

$$(18'') \quad \begin{cases} | A(a + u, b + v)x - A(a, b)x | \leqq M(| u | + | v |) \cdot | x |, \\ | B(a + u, b + v)y - B(a, b)y | \leqq M(| u | + | v |) \cdot | y |. \end{cases}$$

Clearly $A(e, e) = B(e, e) = I$, the identical linear transformation—since irrespective of u, $u \circ e = e \circ u = u$.

Once these definitions and this fact have been stated, the proof of Theorem 8 follows familiar lines. Assuming the existence everywhere and continuity at $x = y = e$ of $A(x, y)$ and $B(x, y)$, one constructs the real functions

$$\phi(\lambda) = | (\lambda x \circ a) - (\lambda x + a) |,$$
$$\psi(\mu) = | (x \circ a \circ \mu y) - (x + a + \mu y) |.$$

Clearly (18′) implies that the upper right-derivatives of $\phi(\lambda)$ and $\psi(\mu)$ are bounded by $|A(\lambda x, a) - I| \cdot |x|$ and $|B(x \circ a, \lambda y) - I| \cdot |y|$, respectively. Hence by the theory of real functions,

$$(3) \quad | (x \circ a \circ y) - (x + a + y) | \leqq K(| a | + | x | + | y |) \cdot (| x | + | y |),$$

where $K(|a| + |x| + |y|)$ is small as long as $|A(\lambda x, a) - I|$ and $|B(x \circ a, \mu y) - I|$ are small identically on $0 \leqq \lambda, \mu \leqq 1$—and so by the continuity of these is an M-function, q.e.d.

Remark. Fréchet's definition obviously specializes to the usual definition of continuous total differentiability when B is finite-dimensional—and is satisfied in this case provided continuous first partial derivatives with respect to all coordinates exist.† (This remark has immediate application to the theory of Lie groups—it shows that if the function $x \circ y = f(x, y)$ has continuous first partial derivatives, then one is dealing with an "analytical group.")

In summary, §§17–19 have contained three alternative definitions of analytical groups, equivalent to Definition 1. One can view these from two angles. They may be regarded from a conceptual angle as giving a better picture of what an analytical group is. Or they may be regarded as giving content to Definition 1 itself—that is, as furnishing examples of analytical groups from other contexts.

Chapter III. Linear groups

20. **Axiomatization.** It is a simple fact, that *one can axiomatize algebras of linear operators‡ on Banach spaces.*

To see this, one must first recall that the operators on *any* linear space B which are defined everywhere, and carry vector sums into vector sums and scalar multiples into scalar multiples, constitute a hypercomplex algebra with

† C. J. de la Vallée-Poussin, *Cours d'Analyse Infinitésimale*, Louvain, 1914, p. 141.

‡ By a "linear operator," we mean ([1], p. 23) any continuous additive, everywhere defined function. This conflicts with the usage for Hilbert spaces, where such operators are called *bounded*.

a principal unit I. (We shall use the notation O for the transformation carrying every $x\epsilon B$ into Θ; I for the identity $x\rightarrow x$; $S, T, U, \cdots$ for other operators.)

One must next observe that if B is a Banach space, then relative to vector sums $T+U$, products λT with scalars, and the "modulus" $\|T\| = \sup_{x\neq\Theta}|Tx|/|x|$ (cf. [1], p. 54), the linear operators on B constitute another Banach space. The proof of this will be left to the reader.†

Finally, $\|T \circ U\| \leqq \|T\| \cdot \|U\|$.

More generally, *any* algebra of linear operators on a Banach space B which contains I and is topologically closed (under the "uniform" topology defined by the metric $\|T-U\|$) has all of the properties just described.

But conversely, let $\mathfrak{H}$ be any system having these properties—i.e., any "metric hypercomplex algebra."‡ Then (applying a classical construction) each element $T\epsilon\mathfrak{H}$ induces a linear transformation θ_T: $X\rightarrow XT$ on the elements $X\epsilon\mathfrak{H}$. Moreover since $\|IT\| = \|I\| \cdot \|T\|$ and $\|XT\| \leqq \|X\| \cdot \|T\|$, the "modulus" of θ_T is precisely $\|T\|$. Thus $\mathfrak{H}$ can be realized as a closed algebra of linear operators on itself, including the identity θ_I: $X\rightarrow XI = X$.

21. **Linear operators with inverses.** Linear operators do not constitute a group under multiplication. But the linear operators S with inverses S^{-1} satisfying $SS^{-1} = S^{-1}S = I$ do. And one can easily prove

Theorem 9. *Let $\mathfrak{H}$ be any metric hypercomplex algebra. Then the map $(I+T)\rightarrow T$ of the elements $(I+T)\epsilon\mathfrak{H}$ with $\|T\| < \frac{1}{2}$ onto the linear space defined by $\mathfrak{H}$, exhibits these elements as an analytical group $\mathfrak{G}$ under multiplication.*

Proof. Refer to Definition 1. The only properties in any doubt are (2′)–(2″). But

$$\begin{aligned} \Xi &\equiv \|[(I+X)\circ(I+Y) - (I+X)\circ(I+Z)] - [Y-Z]\| \\ &= \|[X\circ Y - X\circ Z]\| = \|X\circ(Y-Z)\|, \quad \text{by algebra}, \\ &\leqq \|X\|\cdot\|Y-Z\|, \qquad\qquad \text{by hypothesis}, \end{aligned}$$

proving (2′). One obtains (2″) similarly.

Theorem 10. *The canonical transformation of $\mathfrak{G}$ is given explicitly by the convergent power series*

$$T \leftarrow \exp(T) - I \equiv T + \frac{1}{2!}T^2 + \frac{1}{3!}T^3 + \cdots = t(P_T^*).$$

† For instance, if $\{T_n\}$ is a fundamental sequence of linear operators, then for any x, $\{T_n x\}$ is a fundamental sequence in B, whose limit we shall define as Tx. By continuity, $T(x+y) = Tx+Ty$ and $|Tx-Ty| \leqq \lim_{n\to\infty}\|T_n\|\cdot|x-y|$.

‡ More properly, any metric *associative* hypercomplex algebra. Omitting the associative law, we get a more general definition (cf. §30), which however yields no realization theorem.

Proof. The questions of convergence are settled by the inequalities $\|T^n\| \leqq \|T\|^n$ and $\|T+U\| \leqq \|T\|+\|U\|$, and the assumption $\|T\| < \frac{1}{2}$. But now dividing P_T into n equal parts, we get by the binomial expansion

$$t(P_T)_\pi^* = \left(I + \frac{1}{n} T\right)^n - I = T + \frac{1}{n^2} C_{n,2} T^2 + \frac{1}{n^3} C_{n,3} T^3 + \cdots$$

which converges (cf. supra) to $\exp(T) - I$.

The inverse of the canonical transformation is of course given by the power series

$$T \leftarrow \log (I + T) = T - \tfrac{1}{2}T^2 + \tfrac{1}{3}T^3 - \cdots,$$

but we shall not use this fact.†

22. **Generalization of Theorem** 9. A metric algebra $\mathfrak{X}$ need not possess a unit I nor satisfy $\|I\| = 1$ in order that the symbolic elements $I + X$ with $X \epsilon \mathfrak{X}$ and $\|X\| < \frac{1}{2}$ should form an analytical group nucleus when multiplied according to the rule

$$(I + X) \circ (I + Y) = I + (X + Y + X \circ Y). \tag{22.1}$$

The arguments of §21 do not involve these assumptions.

An important example of such an algebra is due to Delsarte [6], and is also cited by Yosida (op. cit.). It is the algebra $\mathfrak{A}$ of all infinite matrices $A = \|a_{ij}\|$ for which $\sum_{i,j} |a_{ij}|^2 < +\infty$. If we set $\|A\|^2 = \sum_{i,j} |a_{ij}|^2$, we have a Banach space, in which products $C = A \circ B = \|c_{ij}\|$ can be defined by the usual rule $c_{ij} = a_{ik}b_{kj}$ —the series being convergent by Schwarz' inequality, and satisfying, besides, $\|C\| \leqq \|A\| \cdot \|B\|$.

The algebra $\mathfrak{A}$ corresponds of course to the algebra of Schmidt kernels in the theory of integral equations, and is isometric with Hilbert space.

23. **Function of composition.** The formulas of the preceding section lead directly to explicit expressions for the function $X \circ Y$ of composition.

Under the original parameters, $X \circ Y = F(X, Y)$ is analytic since (by the distributivity of multiplication) it is *linear in both variables*—and among the functions between linear spaces, next to constant functions, linear functions

† Remark: The above treatment was suggested by that of J. von Neumann [12]. The main changes are: explicit discussion of transformations as abstract elements, and use (following Banach) of the "modulus" $\|X\|$ for norm.

The concept of a metric hypercomplex algebra ("complete normed vector ring") was announced by the author in Abstract 41-3-104 of the Bulletin of the American Mathematical Society (1935); a similar definition is given by K. Yosida (*On the group embedded in the metrically complete ring*, Japanese Journal of Mathematics, vol. 13 (1936), pp. 7–26). Yosida does not require $\|I\| = 1$; cf. §22.

Another example, discussed at length by M. H. Stone, consists of the linear operators T_a: $f(x) \rightarrow f(x)a(x)$ on the space of bounded functions on an abstract class. This is a closed subalgebra of the algebra of §20.

are the most purely analytic. Thus the second partial derivatives of $F(X, Y)$ are constant and so the higher derivatives all vanish identically.

Moreover by Theorem 10, if we denote by X^m as usual $X \circ \cdots \circ X$, then we have the following explicit expression for $X \circ Y = G(X, Y)$ *under canonical parameters*,

$$\begin{aligned} G(X, Y) &= \log (I + F(\exp X - I, \exp Y - I)) \\ &= \log\left(I + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{1}{m!n!} F(X^m, Y^n)\right) \\ &= \sum_{k=1}^{\infty} (-1)^{k-1}/k \left\{\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{1}{m!n!} F(X^m, Y^n)\right\}^k \end{aligned}$$

whose first terms can be found easily, and are monomials.

It has been shown by J. E. Campbell [3] and F. Hausdorff [8] that this series can also be developed in terms of X, Y, and iterations of the bilinear function $[X, Y] = XY - YX$. The resulting "SCH-series" will be proved in Chapter V to be valid also for non-linear analytical groups.

24. **Digression: polynomials and analyticity.** The *algebraic* significance of the SCH-series will be discussed in Chapter V; what about its *analytical* significance?

It exhibits $G(X, Y)$ as analytical in the strong sense that (1) it is the limit of an absolutely convergent series of polynomials of increasing degrees,† (2) its derivatives all exist and can be found through term-by-term differentiation of the series, (3) hence the Taylor's series for $G(X, Y)$ converges absolutely to $G(X, Y)$—all within a sphere of positive radius.

Although it is not entirely clear when a function between Banach spaces is "analytical"—there may be various generalizations of the established notion for functions between euclidean spaces—it seems undeniable that at least any function with properties (1)–(3) should be called analytical.

25. **Adjoint of an analytical group.** In the present section, we shall show that the notion of the *adjoint* of a Lie group can be extended without real modification to the case of analytical groups. We state this more precisely in the following theorem:

† For polynomial functions between Banach spaces, cf. S. Mazur and W. Orlicz, *Grundlegende Eigenschaften der polynomische Operatoren*, Studia Mathematica, vol. 5 (1935), pp. 50–68 and pp. 179–189. One can define polynomials through continuity + the identical vanishing of $(n+1)$st differences, through the identical vanishing of $(n+1)$st derivatives, or as sums of multilinear functions in a variable repeated $0, \cdots, n$ times; and these definitions are equivalent.

Unlike these authors, we are concerned with functions of two variables. N.B.: A polynomial function on r variables which is homogeneous of degree k in each, is homogeneous of degree kr (and not of degree k) on the product-space of the variables.

THEOREM 11. *Let G be any analytical group under canonical parameters. Then each element $g\epsilon G$ determines a linear transformation θ_g: $x\to g^{-1}xg$ on the parameter-space of G, and the correspondence $g\to\theta_g$ is continuously homomorphic.*

Proof. Since G is a topological group θ_g is a topological automorphism. Hence by Corollary 6.1 it is a linear transformation on G. Moreover by the well-known identity $(gh)^{-1}x(gh)=h^{-1}(g^{-1}xg)h$, the correspondence $g\to\theta_g$ is homomorphic. It remains only to show that it is continuous under the *uniform* topology. But

$$\begin{aligned} |h^{-1}xh - g^{-1}xg| &= O(|h^{-1}x^{-1}hg^{-1}xg|), \quad \text{by (3)}, \\ &= O(|g^{-1}[(hg^{-1})^{-1}x^{-1}(hg^{-1})x]g) \\ &= O(|(hg^{-1})^{-1}x^{-1}(hg^{-1})x) \\ &\qquad (\text{since } \|\theta_g\| = O(|g|), \quad \text{by } (14\beta)) \\ &= O(|g-h|)\cdot|x|, \quad \text{by } (27\beta), \end{aligned}$$

whence $\|\theta_g-\theta_h\|=O(|g-h|)$, completing the proof.

The validity of the proof of course depends on proving (27β) without the aid of Theorem 11. We shall do this in §27.

The correspondence $g\to\theta_g$ does not always carry open sets into open sets: it need not be "gebietstreu" in the sense of Freudenthal.

CHAPTER IV. COMMUTATION

26. **Outline.** The present chapter will be devoted to showing how every analytical group G possesses a bilinear "commutation function" $[x, y]$. In Chapter V, it will be shown that $[x, y]$ determines G to within local isomorphism.

The commutation function $[x, y]$ belonging to a given group G is most easily defined as the bilinear asymptote at $x=y=0$ to the purely algebraic commutator function

$$(x, y) = K(x, y) \equiv x^{-1}y^{-1}xy.$$

The fact that (x, y) *has* a bilinear asymptote is proved below (in §28) from the relations (deduced in §27)

$$(27\alpha) \qquad \begin{cases} |(u\circ x, y) - (x, y) - (u, y)| \leqq M(|x|+|y|)\cdot|(u, y)|, \\ |(x, v\circ y) - (x, y) - (x, v)| = M(|x|+|y|)\cdot|(x, v)|, \end{cases}$$

$$(27\beta) \qquad |(x, y)| = O(|x|\cdot|y|),$$

while the fact that $[x, y]$ is a topologico-algebraic invariant associated with G is almost obvious (cf. §29).

Moreover one can deduce the familiar identities

$$(30\alpha)\qquad \begin{cases} [x, y] + [y, x] = 0, \\ [[x, y], z] + [[y, z], x] + [[z, x], y] = 0 \end{cases}$$

as corollaries of formal identities on group products. These results can be summarized in the statement that G possesses a *metric Lie algebra* $L(G)$. Chapter IV concludes with various applications of $L(G)$, to the case that G is under canonical parameters.

In the proofs of Chapter IV, group algebra plays a novel and essential role.

27. **The approximate bilinearity of** $K(x, y)$. The present section will be devoted to showing that $(x, y) = K(x, y)$ is approximately bilinear at $x = y = 0$, in the sense that (27α)–(27β) are true.

The proof of (27α) is almost immediate. One has the formal identity

$$(27\gamma)\qquad \begin{aligned} (u \circ x, y) &= x^{-1}u^{-1}y^{-1}uxy \\ &= (u, y)_x \circ (x, y) \end{aligned}$$

under the convention that g_x denotes $x^{-1}gx$. But by the fundamental inequality (2) of §8,

$$(27\delta)\qquad |(u, y)_x - (u, y)| \leqq M(|x|) \cdot |(u, y)|,$$

whence $|(u, y)_x| = O(|(u, y)|)$. Hence

$$\begin{aligned} \Xi &\equiv |(u \circ x, y) - (x, y) - (u, y)| \\ &\leqq |(u \circ x, y) - (x, y) - (u, y)_x| + |(u, y)_x - (u, y)| \\ &\leqq M(|(x, y)|) \cdot |(u, y)_x| + M(|x|) \cdot |(u, y)| \\ &\qquad\qquad\qquad \text{(by } (27\gamma) \text{ and (3) of §9, and } (27\delta)) \\ &\leqq M(|x| + |y|) \cdot |(u, y)|, \end{aligned}$$

since $|(u, y)_x| = O(|(u, y)|)$. But this is the first half of (27α); the second half follows from the symmetry between left- and right-multiplication.

As a special instance of (27α), we have

$$|(x^m, y) - (x^{m-1}, y) - (x, y)| \leqq M(|x^m| + |y|) \cdot |(x, y)|.$$

Hence, since $|x^m| = O(|mx|) = O(m|x|)$, by induction

$$|(x^m, y) - m(x, y)| \leqq M(m|x| + |y|) \cdot m \cdot |(x, y)|.$$

Combining with the symmetric formula in (x, y^n), we get

$$(27\epsilon)\qquad |(x^m, y^n) - mn(x, y)| = M(m|x| + n|y|) \cdot mn |(x, y)|.$$

Consequently, within some small radius ρ of the origin,

$$(27\zeta) \qquad |(x, y)| \leqq O(|(x^m, y^n)|/mn).$$

But clearly within this sphere, given $x\neq 0$ and $y\neq 0$, one can so choose m and n that $\frac{1}{2}\rho<|x^m|, |y^n|<\rho$—whence, $|(mx, ny)|$ being bounded within this sphere, we get $|(x^m, y^n)| \leqq O(|mx|\cdot|ny|)$, and so by (27ζ),

$$(27\beta) \qquad |(x, y)| \leqq O(|x|\cdot|y|).$$

It is a corollary, since $y\circ x\circ K(x, y)=x\circ y$, and likewise $(y\circ x)+(x\circ y-y\circ x)=x\circ y$, that (by (3))

$$(27\beta') \qquad |xy-yx| \leqq O(|x|\cdot|y|).$$

28. **The asymptote** $[x, y]$. Substituting from (27β) in (27α), and recalling that $x+w=u\circ x$ implies† $|w|\sim|u|$, we get

$$(28\alpha) \qquad \begin{cases} |(x+w, y)-(x, y)-(w, y)| \leqq M(|x|+|y|)\cdot|w|\cdot|y|, \\ |(x, y+w)-(x, y)-(x, w)| = M(|x|+|y|)\cdot|x|\cdot|w|, \end{cases}$$

from which there follows

$$(28\alpha') \qquad |(x+u, y+v)-(x, y)| \leqq O(|u|+|v|).$$

Now start anew with (28α)–$(28\alpha')$, and use the same algebraic analysis used in proving (27ϵ). By (28α),

$$|(mx, y)-((m-1)x, y)-(x, y)| \leqq M(m|x|+|y|)\cdot|x|\cdot|y|.$$

Hence by induction on m, we get

$$|(mx, y)-m(x, y)| \leqq M(m|x|+|y|)\cdot m|x|\cdot|y|.$$

Combining with the symmetric formula in (x, ny), we have

$$(28\beta) \qquad |(mx, ny)-mn(x, y)| \leqq M(m|x|+n|y|)\cdot mn\cdot|x|\cdot|y|.$$

By double use of (28β), we get for $0<h/m, k/n<1$,

$$\left|\left(\frac{h}{m}x, \frac{k}{n}y\right)-\frac{hk}{mn}(x, y)\right| \leqq M(|x|+|y|)\cdot\frac{hk}{mn}\cdot|x|\cdot|y|,$$

whence, by rational approximation and passage to the limit, using $(28\alpha')$ to establish continuity, we have

$$(28\gamma) \qquad \left|\frac{1}{\lambda\mu}(\lambda x, \mu y)-(x, y)\right| \leqq M(|x|+|y|)\cdot|x|\cdot|y|$$

for $0<\lambda, \mu<1$. Therefore if $\lambda+\mu+\lambda'+\mu'<\epsilon$, then

† By $|w|\sim|u|$ we mean that $\big||w|-|u|\big| \leqq M(|x|+|u|+|w|)\cdot|u|$; this relation is evidently reflexive, symmetric, and transitive.

$$\left|\frac{1}{\lambda\mu}(\lambda x, \mu y) - \frac{1}{\lambda'\mu'}(\lambda' x, \mu' y)\right| = M(\epsilon)\cdot |x| \cdot |y|$$

and so, by the completeness of the parameter-space

$$[x, y] \equiv \lim_{\lambda,\mu\downarrow 0} \frac{1}{\lambda\mu} K(\lambda x, \mu y) \tag{28δ}$$

exists. Furthermore, by (28γ),

$$|[x, y] - (x, y)| \leqq M(|x| + |y|)\cdot |x| \cdot |y|. \tag{28ε}$$

Finally, since $(-x+x, y) = (0, y) = 0$, by (28α)

$$|(-x, y) - [-(x, y)]| \leqq M(|x| + |y|)\cdot |x| \cdot |y|.$$

whence we see that

$$[x, y] = \lim_{\lambda,\mu\to 0} \frac{1}{\lambda\mu}(\lambda x, \mu y) \tag{28ζ}$$

exists.

29. **The bilinearity of $[x, y]$, etc.** In this section, we shall prove the bilinearity and topologico-algebraic invariance of $[x, y]$.

The invariance of $[x, y]$ under continuous isomorphisms between groups under canonical parameters follows from the definition and Theorem 6. And by (28α), "distortions" of type (11a) change $(\lambda x, \mu y)$ by $o(|\lambda|\cdot|\mu|)$, from which invariance under general continuous isomorphisms follows by Corollary 6.2.

As for the bilinearity of $[x, y]$, by (28α)

$$\Xi_1 \equiv \left|\frac{1}{\lambda\mu}(\lambda[x+u], \mu y) - \frac{1}{\lambda\mu}(\lambda x, \mu y) - \frac{1}{\lambda\mu}(\lambda u, \mu y)\right|$$
$$= M(|\lambda x| + |\mu y|)\cdot\frac{1}{\lambda\mu}\cdot |\lambda u| \cdot |\mu y|$$

whence, passing to the limit, $[x+u, y] - [x, y] - [u, y] = 0$. Hence $[x+u, y] = [x, y] + [u, y]$; and $[x, y+v] = [x, y] + [x, v]$ by symmetry. Also, by (27β) and (28ε), $[x, y] = O(|x|\cdot|y|)$, and so is bounded. Hence it is bilinear.

In summary of the above results,

THEOREM 12. *(x, y) has a bilinear asymptote $[x, y]$ which is a topologico-algebraic invariant of G.*

Remark. In a linear group, algebra based on the expansion $(I+\lambda X)^{-1} = I - \lambda X + \lambda^2 X^2 - \lambda^3 X^3 + \cdots$ shows

$$\frac{1}{\lambda^2}(\lambda X, \lambda Y) = (XY - YX) + \text{terms of higher order}$$

whence, passing to the limit, $[X, Y] = XY - YX$.

30. **Metric Lie algebras.** One can now deduce relations (30α) from algebraic identities on group products.

In the first place, since $(y, x) = (x, y)^{-1}$, and $u^{-1} + u$ is nearly zero, clearly $[x, y] + [y, x] = 0$. That is, $[x, y]$ is skew-symmetric.

The proof that $[x, y]$ satisfies Jacobi's identity,

$$[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$$

is less simple. It depends very essentially on realizing that by (27β) and (28ϵ),

$$\begin{aligned} \Xi_2 &\equiv |((x, y), z) - [[x, y], z]| \\ &\leqq |((x, y), z) - ([x, y], z)| + |([x, y], z) - [[x, y], z]| \\ &\leqq M(|x| + |y| + |z|) \cdot |x| \cdot |y| \cdot |z| \end{aligned} \tag{30β}$$

and, besides, on remarking that since $v \circ u = u \circ v \circ (v, u)$, to permute two commutators in a group product changes the value by an amount which is by (27β) small to the fourth order.

But direct computation based on cancellation proves†

$$(x, y)((x, y), z)(z, y)(z, x)((z, x), y)(y, x)(y, z)((y, z), x)(x, z) = \Theta.$$

Therefore, permuting terms, and cancelling

$$(x, y)(y, x) = (z, y)(y, z) = (z, x)(x, z) = \Theta,$$

we get by the preceding remark, the inequality

$$|((x, y), z)((y, z), x)((x, z), y)| \leqq O(|x| + |y| + |z|) \cdot |x| \cdot |y| \cdot |z|.$$

Hence by (30β) and the fundamental relation (3),

$$|[[x, y], z] + [[y, z], x] + [[z, x], y]| = M(|x| + |y| + |z|) \cdot |x| \cdot |y| \cdot |z|.$$

Replacing x, y, z by $\lambda x, \lambda y, \lambda z$ where λ is small, and using linearity, we get Jacobi's Identity in the limit.

Summarizing, we may say (in the language of Chapter III),

Theorem 13. *Relative to sums $x+y$, scalar products λx, and "brackets" $[x, y]$, the parameter-space of any analytical group nucleus G is a metric Lie algebra $L(G)$.*

Remark 1. In §§26–30 we have nowhere assumed that G was under canonical parameters.

† This formula was suggested to the author by identities in §2.3 of [7].

Remark 2. Since $|[x, y]| \leqq O(|x| \cdot |y|)$, after changing the scale (i.e., multiplying the norm by a suitable constant) we can assume simply $|[x, y]| \leqq |x| \cdot |y|$.

Remark 3. Brackets $[x, y]$ are defined for all x, y in the Banach space B, unlike $x \circ y$ which is defined only locally.

We shall show (Corollary 15.1) that G is determined to within local isomorphism by $L(G)$, and that conversely any metric Lie algebra belongs to an analytical group nucleus. This shows that the problem of enumerating the analytical group nuclei with a given parameter-space is equivalent to that of enumerating the different metric Lie algebras on the same linear space.

31. **Subgroups and normal subgroups.** The results of §§31–32 will refer to analytical groups G under canonical parameters. A subset S of elements of G will be called an *analytical subgroup nucleus* if and only if, relative to the topology and group multiplication table of G, S is itself an analytical group nucleus. An analytical subgroup nucleus S will be called *normal* (or invariant) if and only if for every $g \epsilon G$, $g^{-1}Sg$ contains some neighborhood of the identity of S.

If S is an analytical subgroup nucleus, then each $x \epsilon S$ must lie on a one-parameter subgroup x^λ in S, and hence (by Theorem 6) a segment λx in G must lie in S. Again, the length of this segment must exceed some fixed positive constant; otherwise we could find $\{x_n\}$ such that $\lambda_n x_n \epsilon S$ implies $\lambda_n x_n \to 0$, and this is impossible in an analytical group nucleus.

Therefore S must contain with x and y, $k(\lambda x, \lambda y)/\lambda^2$ for some fixed $k > 0$ and all λ on $[0, 1]$; hence it must contain with x and y, $k[x, y]$ (since, being *complete*, it is *closed* in G). Similarly, it must contain with x and y, $x + y = \lim_{\lambda \to 0}(\lambda x \circ \lambda y)/\lambda$. And finally, if two such subgroup nuclei contain elements on the same class of segments $\lambda x \epsilon G$, then they clearly generate (in case G is a group) the same subgroup of G, and so may be identified.

These facts may be summarized in

(31α) *Let G be any analytical group nucleus under canonical parameters. The analytical subgroup nuclei S of G are pieces of closed subalgebras of the metric Lie algebra $L(G)$, two subgroups being identical if and only if the subalgebras are.*

If S is "normal" (i.e., invariant under all inner automorphisms), then $x \epsilon S$ and $g \epsilon G$ imply that for some $k > 0$, $k[g, x] = \lim_{\lambda \to 0}(k\{\lambda x \circ \lambda g_{\lambda x}\}/\lambda^2) \epsilon S$. Furthermore,

$$g + x = \lim_{n \to \infty} \left[\left(\frac{1}{n}\, g\right) \circ \left(\frac{1}{n}\, x\right)\right]^n$$

$$= g \circ \prod_{k=1}^{n} \left\{ \left(\frac{1}{n} g\right)^{k-n} \circ \left(\frac{1}{n} x\right) \circ \left(\frac{1}{n} g\right)^{n-k} \right\},$$

$$\epsilon gS = Sg.$$

Hence $g+S=Sg$, and

(31β) *If S is normal, then the associated subalgebra of $L(G)$ is invariant*† *and the cosets of S are the hyperplanes parallel to the manifold of S.*

We shall prove converses to (31α)–(31β) in (32β) and corollary 15.5.

32. **The adjoint group.** Using the commutation function, one can easily deduce an explicit series for the adjoint group of §25.

Define T: $u \rightarrow T(u) = (y/n)^{-1} \circ u \circ (y/n)$. Then T is a linear transformation, and

$$\left| T(u) - \left(u + \left[u, \frac{1}{n} y\right]\right) \right| \leqq \left| u \circ \left(u, \frac{1}{n} y\right) - u - \left(u, \frac{1}{n} y\right) \right|$$

$$+ \left| \left(u, \frac{1}{n} y\right) - \left[u, \frac{1}{n} y\right] \right|$$

$$\leqq \frac{1}{n} \cdot M\left(|u| + \frac{1}{n} |y| \right) |u| \cdot |y|$$

(by (3), (27β) and (28ϵ)).

Hence by n-fold iteration and the binomial expansion,

$$\left| T^n(u) - \left\{ u + [u, y] + C_{n,2} \cdot \frac{1}{n^2} [[u, y], y] + \cdots \right\} \right|$$

$$\leqq M\left(|u| + \frac{1}{n} |y| \right) |u| |y|.$$

Whence, since $T^n(u) = y^{-1} \circ u \circ y$, passing to the limit, we have

$$|w(u, y)| \equiv \left| \{y^{-1} \circ u \circ y\} - \left\{ u + [u, y] + \frac{1}{2!} [[u, y], y] + \cdots \right\} \right|$$

$$= o(|u|) \cdot |y|.$$

But since the terms are all linear in u, clearly

$$|w(u, y)| = n \left| w\left(\frac{1}{n} u, y\right) \right| = n \cdot o(1/n) \cdot |u| \cdot |y|.$$

That is, letting $n \uparrow \infty$, $|w(u, y)| = 0$, and so

† In the usual sense, that $x \epsilon S$ and $g \epsilon L(G)$ imply $[g, x] \epsilon S$.

(32α) $y^{-1}xy = x + [x, y] + \frac{1}{2!}[[x, y], y] + \frac{1}{3!}[[[x, y], y], y] + \cdots.$

From (32α) we deduce as a corollary,

(32β) *If the subalgebra associated with a given subgroup S of an analytical group G is invariant, then S is a normal subgroup.* (Converse of (31β).)

Chapter V. Function of composition

33. **Introduction.** The main purpose of this chapter is to show in §§34–36 how the function $x \circ y = f(x, y)$ of composition of any analytical group under canonical parameters, can be written as the sum of an infinite series of polynomials determined by the commutation† function $[x, y]$—and to deduce in §37 various corollaries from this fact.

F. Schur [16] first showed that this series was valid in all groups under canonical parameters. Campbell [3] and Hausdorff [8] have since obtained it by other methods,‡ and so we shall call it the "*SCH*-series."

The present exposition is preferable on three grounds to those cited. It applies to infinite-dimensional groups. It paraphrases identities on pure group products [§§38–40], and does not require Taylor's series or manipulations with matrix polynomials (which are unnatural in non-linear groups). And most important, it generalizes easily to yield similar series expressing the definite (product) integrals over fixed time intervals of variable linear combinations of infinitesimal transformations, in a form which (like the *SCH*-series§) is independent of the group which they generate.

In §§38–40, the paraphrases in terms of identities on group products, of the *SCH*-series and other identities in the theory of continuous groups, are developed. They are not a part of the technical argument—unlike the paraphrases of the identities of Lie-Jacobi, which are actually used in proving the latter. They have been included because they correlate the theories of discrete groups and continuous groups in a way essential to the full understanding of either.

---

† Expressions (x, y) or $[x, y]$ will be called "simple" commutators and brackets, respectively; the commutator (ϕ, ψ) of any two commutators ϕ and ψ of "lengths" $w(\phi)$ and $w(\psi)$—where for uniformity individual letters are regarded as commutators of length one—will be called a "complex" commutator of "length" $w(\phi)+w(\psi)$. Similarly with complex brackets $[\phi, \psi]$ of "length" $w(\phi)+w(\psi)$.

‡ Schur starts with the obvious identity $f(x, (\lambda+\delta)y) = f(x, \lambda y) \circ \delta y$, determines $d/d\lambda\{f(x, \lambda y)\}$ from this, and integrates the resulting differential equation. Campbell and Hausdorff develop the series by setting $e^x e^y = e^{f(x,y)}$, and use the algebra of matrices to solve for $f(x, y) = \log[1+(e^x e^y - 1)]$—thus introducing an extraneous operation of addition.

§ The *SCH*-series is the case where x operates first for a unit of time, followed by y operating for a unit of time.

34. **Product-equivalence of paths.** Let G be any analytical group under canonical parameters. Then

(34α) *The problem of determining* $x \circ y = f(x, y)$ *is equivalent to that of determining, given two short paths* P *and* Q, *whether or not* $t(P^*) = t(Q^*)$.

(We shall express the relation $t(P^*) = t(Q^*)$ by writing $P \sim Q$, and saying P is *product-equivalent* to Q. By (14θ), $P \approx Q$ implies $P \sim Q$.)

Proof. Since $x \circ y = t((P_x \oplus P_y)^*)$ under canonical parameters, we have found the $z = f(x, y)$ when‡ we have found the $P_z \sim P_x \oplus P_y$. While conversely, $t(Q^*)$ is approximated arbitrarily closely and hence determined by the $t(Q_\pi^*) = t(Q_1) \circ \cdots \circ t(Q_r)$ for the different partitions π of Q—and the $t(Q_1) \circ \cdots \circ t(Q_r)$ are determined by Q and the function of composition.

From (34α) and the known existence of an SCH-series expressing $f(x, y)$ in terms of the commutation function, we certainly can infer that $t(Q^*)$ is determined by Q and the commutation function in a way valid in all groups G under canonical parameters. But it by no means gives us explicit series for Q^* (except when $Q = P_x \oplus P_y$)—and it is such series that we shall finally obtain, getting the SCH-series as a special case (cf. §36).

Our first step will be to determine, given Q, all the $P \sim Q$. To this end we prove

(34β) *Let* P *and* Q *be any admissible paths with domain* $[0, \Lambda]$. *Then* $P \sim Q$ *if and only if some* U: $u(\lambda)$ *exists, such that* $u(0) = u(\Lambda) = 0$ *and*

$$| \delta p - [u^{-1}(\lambda) \circ \delta q \circ u(\lambda) + \delta u\dagger] | \leqq o(| \Delta Q | + | \Delta U |).$$

Proof. Suppose $P \sim Q$, and write $p^*(\lambda) = q^*(\lambda) \circ u(\lambda)$. Since $p^*(0) = 0 = q^*(0)$ and $p^*(\Lambda) = t(P^*) = t(Q^*) = q^*(\Lambda)$, $u(0) = u(\Lambda) = 0$. Define $R = P^*$, so that $P = R\dagger$. Clearly if Δ: $[\lambda, \mu]$ is any interval, then by (14η)

$$\begin{aligned} t((\Delta P)^*) &= t((\Delta R\dagger)^*) = r^{-1}(\lambda) \circ r(\mu) \\ &= \{u^{-1}(\lambda) \circ [q^{*-1}(\lambda) \circ q^*(\mu)] \circ u(\lambda)\} \circ \{u^{-1}(\lambda) \circ u(\mu)\} \\ &= \{u^{-1}(\lambda) \circ t((\Delta Q)^*) \circ u(\lambda)\} \circ t((\Delta U\dagger)^*). \end{aligned}$$

But $|\Delta P| = O(|\Delta R|) \leqq O(|\Delta Q^*| + |\Delta U|) \leqq O(|\Delta Q| + |\Delta U|)$; besides $|t((\Delta P)^*) - t(\Delta P)| \leqq o(|\Delta P|)$, and similarly with ΔQ and ΔU (by (14ϵ)) even after the inner automorphism induced by $u(\lambda)$. Moreover by (3) $|x \circ y - (x+y)| = o(|x| + |y|)$; consequently if we write $\delta p = t(\Delta P)$, $\delta q = t(\Delta Q)$ and $\delta u = t(\Delta U)$, we get

$$| \delta p - [u^{-1}(\lambda) \circ \delta q \circ u(\lambda) + \delta u\dagger] | = o(| \Delta Q | + | \Delta U |).$$

‡ We recall the notation P_x for the path $p_x(\lambda) = \lambda x$ defined on $[0, 1]$, and $P_x \oplus P_y$ for the broken line R: $r(\lambda) = \lambda x$ on $[0, 1]$, and $r(\lambda) = x + (\lambda - 1)y$ on $[1, 2]$.

Conversely, suppose that this inequality is satisfied for some $u(\lambda)$ (of bounded variation‡) with $u(0)=u(\Lambda)=0$. Then, when we write $r(\lambda)=q^*(\lambda) \circ u(\lambda)$, obviously $t((R\dagger)^*)=t(R)=t(Q^*)$. Moreover by the argument above, $r\dagger(\lambda)$ satisfies the given inequality. Therefore by the triangle inequality,

$$|\delta p - \delta r\dagger| \leqq M(|\Delta Q| + |\Delta U|)\cdot(|\Delta Q| + |\Delta U|).$$

Hence if π is any partition of $[0, \lambda]$, writing $\|\pi\|$ for sup $(|\Delta Q|+|\Delta U|)$, and summing inequalities, we get

$$\begin{aligned}|p(\lambda) - r\dagger(\lambda)| &= |[p(\lambda) - p(0)] - [r\dagger(\lambda) - r\dagger(0)]| \\ &\leqq M(\|\pi\|)\cdot(|Q| + |U|),\end{aligned}$$

whence in the limit $p(\lambda)\equiv r\dagger(\lambda)$.

We can rewrite (34β) perhaps more suggestively in the notation of differentials, as

$$(34\gamma) \qquad dp = u^{-1}(\lambda) \circ dq \circ u(\lambda) + du\dagger.$$

35. **Devices for calculation.** Consider the terms of this formula. By (32α), $u^{-1} \circ dq \circ u$ can be calculated explicitly from U, Q and the commutation function.

Again, although we have not shown how to calculate $U\dagger$ from U explicitly§ by using the commutation function, we can now do so in case U is *unidimensional*.

(A path U: $u(\lambda)$ will be called "unidimensional" if and only if it is confined to a straight line—i.e., if and only if for some u_0, $u(\lambda)=\alpha(\lambda)u_0$. If U is unidimensional, then by Theorem 5, $U_\pi^*=U_\pi\dagger=U$ identically, whence in the limit $U^*=U\dagger=U$. By a "unidimensional alteration" of any path Q with domain $[0, \Lambda]$, will be meant any path $P=R\dagger$ determined from an R: $r(\lambda)=q^*(\lambda) \circ [\alpha(\lambda)u_0]$ for which $\alpha(0)=\alpha(\Lambda)=0$. In this case, clearly $P\sim Q$ and furthermore by (34γ),

$$(35\alpha) \qquad dp = u^{-1}(\lambda) \circ dq \circ u(\lambda) + du.$$

And so P is determined by Q, $u(\lambda)$ and the commutation function.)

Since (32α) gives an infinite series in any case, the fact that only unidimensional alterations can be computed explicitly suggests the following procedure: decomposing a given Q into undimensional constituents, altering these one at a time, and justifying the computations by proving general properties of paths represented by infinite series. This we shall do, first proving

‡ I.e., such that the curve U: $u(\lambda)$ is rectifiable.

§ N.B.: $U\dagger$ differs from U by $M(|U|)$—and hence one can deform a given Q little by little into any desired shape (e.g., a straight ray), whose final position will be *determined* by Q and the commutation function. But its *calculation* involves integrating a (highly involved) differential equation.

(35β) *Let* $u_1(\lambda)$, $u_2(\lambda)$, $u_3(\lambda)$, $\cdots$ *be any twice differentiable functions with domain* $[0, \Lambda]$ *and values in a Banach space. Suppose that the* sup $|u_k| \equiv \sup_{0\leq\lambda\leq\Lambda}|u_k(\lambda)|$, *the* sup $|u_k'|$, *and the* sup $|u_k''|$ *all form convergent series. If* $[\lambda, \lambda+d\lambda]$ *is any subinterval of* $[0, \Lambda]$, $u(\lambda)$ *denotes* $\sum_{k=1}^{\infty} u_k(\lambda)$, *and* δu_k *denotes* $u_k(\lambda+d\lambda)-u_k(\lambda)$, *then*

$$\left| \delta u - d\lambda\left[\sum_{k=1}^{\infty} u_k'(\lambda)\right]\right| \leq O(|d\lambda|^2) \leq o(|d\lambda|).$$

Remark. It is a corollary that u is differentiable and has $\sum_{k=1}^{\infty} u_k'(\lambda)$ for derivative.

Proof. Since by the comparison test, all the series involved converge absolutely (and uniformly!), and the terms of absolutely convergent series can be permuted, clearly $\delta u = \sum_{k=1}^{\infty}\delta u_k$. Moreover for every k,

$$|\delta u_k - u_k'(\lambda)d\lambda| \leq \tfrac{1}{2}[\sup |u_k''|]\cdot d\lambda^2.$$

Summing, we get by the triangle law

$$\left| \delta u - d\lambda\left[\sum_{k=1}^{n} u_k'(\lambda)\right]\right| \leq d\lambda \sum_{n+1}^{\infty} \sup |u_k'(\lambda)| + d\lambda^2 \cdot \sum_{k=1}^{n} \sup |u_k''|.$$

When we pass to the limit, this becomes

$$\left| \delta u - d\lambda\left[\sum_{k=1}^{\infty} u_k'(\lambda)\right]\right| \leq d\lambda^2 \cdot \sum_{k=1}^{\infty} \sup |u_k''| \leq O(|d\lambda|^2).$$

It will be convenient to signify that the hypotheses of (35β) are satisfied by writing

$$U = U_1 + U_2 + U_3 + \cdots = \sum_{k=1}^{\infty} U_k.$$

We shall now get a path $R\dagger \sim P_x \oplus P_y$, from which we shall be able to calculate $f(x, y)$ by using an algorithm applicable to all analytical combinations of unidimensional paths. (The analyticity of $P_x \oplus P_y$ is concealed.)

THEOREM 14. *Let* R: $r(\lambda) = \lambda x \circ \lambda y$ *be defined on* $[0, 1]$. *Then* $P_x \oplus P_y \sim R\dagger$. *And (assuming* $|[x, y]| \leq |x| \cdot |y|$ *by* §30) *if* $|x| + |y| < 1/10$, *then*

$$r\dagger(\lambda) = \lambda y + \lambda x + \frac{\lambda^2}{2!}[x, y] + \frac{\lambda^3}{3!}[[x, y], y] + \cdots \equiv s(\lambda).$$

Proof. It is obvious from identities established in §14 that

$$t((P_x \oplus P_y)^*) = t(P_x) \circ t(P_y) = x \circ y = t(R) = t((R\dagger)^*).$$

The proof is complete if we can show that $|\delta r\dagger - \delta s| \leqq o(|d\lambda|)$. For if this is so, then the upper right-derivative of $|r\dagger(\lambda) - s(\lambda)|$ is zero everywhere, and so $r\dagger(\lambda) = s(\lambda)$. But by (35$\beta$), δs differs from $d\lambda\{y+x+\lambda[x, y]+ \cdots\}$ by $o(|d\lambda|)$—and this is by (32α) $d\lambda\{y+y^{-\lambda} \circ x \circ y^{\lambda}\}$. Again, by (3) $d\lambda\{y+y^{-\lambda} \circ x \circ y^{\lambda}\}$ differs from

$$t((\delta R\dagger)^*) = (x^{\lambda} \circ y^{\lambda})^{-1} \circ (x^{\lambda+d\lambda} \circ y^{\lambda+d\lambda})$$
$$= (y^{-\lambda} \circ x^{d\lambda} \circ y^{\lambda}) \circ y^{d\lambda}$$

by $M(|d\lambda| \cdot |x|) \cdot |d\lambda| \cdot |y| \leqq o(|d\lambda|)$. And by (14$\epsilon$) we have $|t((\delta R\dagger)^*) - \delta r\dagger| \leqq o(|d\lambda|)$—completing the chain of links of length $o(|d\lambda|)$ between δs and δr, and hence the proof.

36. **Evaluation of regular paths.** We can now find $f(x, y) = t(R) = t((R\dagger)^*)$ by a process which enables one to find series expressing $t(P^*)$ for *any* short path P which is "regular" in a sense defined below.

Accordingly, let G be any analytical group under canonical parameters, in which a scale of length has been so chosen that $[x, y] \leqq |x| \cdot |y|$. Let P be any path in G which can be written

$$P = P_1 + P_2 + P_3 + \cdots \qquad \text{(in the sense of (35}\beta)),$$

$$P_i\colon\ p_i(\lambda) = \rho_i(\lambda) \cdot b_i \qquad (0 \leqq \lambda \leqq 1),$$

where (1) the $\rho_i(\lambda)$ are analytical scalar functions with $\sum_{i=1}^{\infty} \int |d\rho_i| < 1/10$, (2) the b_i are brackets in elements $x_1, \cdots, x_r$ arranged in order of increasing length, and containing with any b_i and b_j, also $[b_i, b_j] = b_{f(i,j)}$. Such a path will be called *regular*.

Remark 1. By inserting dummy terms $0 \cdot b_i$, one can make *any* sum of scalar multiples of brackets in $x_1, \cdots, x_r$ satisfy (2) simply because the number of different brackets of any preassigned length w in $x_1, \cdots, x_r$, is finite.

Remark 2. If $|x| + |y|$ is small enough, then the $r\dagger(\lambda)$ of Theorem 14 is regular.

Remark 3. Since $|[x, y]| \leqq |x| \cdot |y|$, $|b_i| \leqq 1$ identically if $|x_1| \leqq 1, \cdots, |x_r| \leqq 1$.

THEOREM 15. *Let P be any regular path. Then $t(P^*)$ is $\sum_{i=1}^{\infty} \gamma_i b_i$, where each γ_i can be calculated from $\rho_1(\lambda), \cdots, \rho_i(\lambda)$ in a finite number of rational operations, integrations, and differentiations. The calculations are independent of G.*

Outline of proof. We shall construct paths $P'' \sim P' = P$, $P''' \sim P''$, $P^{\text{iv}} \sim P'''$, $\cdots$ by successive unidimensional alterations. Each $P^{\nu+1}$ will be "regular" in the same sense that P is, except that $1/10$ may be replaced by some other constant $< 1/5$. Moreover the $\rho_i^{\nu+1}(\lambda)$ for $i \leqq \nu$ will be of the

form $\lambda\gamma_i$—where γ_i is independent of ν—and the $\rho_i^{\nu+1}(\lambda)$ for $i>\nu$ will be increasingly negligible—whence $t(P^*)=\sum_{i=1}^{\infty}\gamma_i\cdot b_i$.

Definition of $P^{\nu+1}$ by induction. If one sets

$$u_\nu(\lambda) = [\lambda\rho_\nu^\nu(1) - \rho_\nu^\nu(\lambda)]\cdot b_\nu \equiv \beta_\nu(\lambda)\cdot b_\nu$$

and can obtain a $P^{\nu+1}=\sum_{i=1}^{\infty}\rho_i^{\nu+1}(\lambda)\cdot b_i$ from P^ν through unidimensional alteration by $u_\nu(\lambda)$, then assuming the term-by-term differentiability of all series, by (32α) and (35α), we obtain heuristically

$$(*)\qquad dp^{\nu+1} = dp^\nu + d\beta_\nu(\lambda)b_\nu + \sum_{j,k=1}^{\infty} d\rho_j^\nu\cdot[\beta_\nu(\lambda)]^k\cdot b_{i(j,k)}\cdot\frac{1}{k!},$$

where $b_{i(j,1)}\equiv[b_j,\ b_\nu]$ and $b_{i(j,k)}\equiv[b_{i(j,k-1)},\ b_\nu]$. But clearly $i(j,\ k)=i$ has in no case an infinity of solutions $(j,\ k)$. Hence we can certainly *define*

$$\rho_i^{\nu+1}(\lambda) = \rho_i^\nu(\lambda) + \sum_{i(j,k)=i}\int\frac{1}{k!}[\beta_\nu(\lambda)]^k\cdot d\rho_j^\nu$$

with the assurance of obtaining analytical $\rho_i^{\nu+1}(\lambda)$—and using only rational operations, integration, and differentiation.

Actual proof. Let us do this. Then—since the length of no $b_{i(j,k)}$ exceeds that of b_ν—certainly by construction $\rho_\nu^{\nu+1}(\lambda)=\lambda\rho_\nu^\nu(1)=\lambda\gamma_\nu$, and for $i<\nu$, $\rho_i^{\nu+1}(\lambda)=\rho_i^\nu(\lambda)=\lambda\gamma_i$ by induction. Furthermore

(36α) *The series* (*) *converge in the sense of* (35β). *Consequently* (*collecting terms*) $P^{\nu+1}=\sum_{i=1}^{\infty}P_i^{\nu+1}$ *in the same sense. Moreover* $\sum_{i=1}^{\infty}\int|d\rho_i^{\nu+1}|<1/5$.

Remark. They even converge absolutely if we replace each bracket by the product of the absolute values of its entries.

Proof. If $\sigma(\lambda)$, $\beta(\lambda)$ and $\rho(\lambda)$ are *any* real analytical functions, then certainly

$$\left\{\begin{aligned}
\sup\ |\sigma| &\leqq \int|d\sigma| = \int|\beta|^k\cdot|d\rho| \\
&\leqq \left[\int|d\beta|\right]^k\cdot\left[\int|d\rho|\right], \\
\sup\ |\sigma'| &\leqq \left[\int|d\beta|\right]^k\cdot\sup\ |\rho'|, \\
\sup\ |\sigma''| &\leqq k\cdot\left[\int|d\beta|\right]^{k-1}\cdot\sup\ |\rho'| + \left[\int|d\beta|\right]^k\cdot\sup\ |\rho''|
\end{aligned}\right.$$

(differentiation is indicated by superscribing primes). Hence by induction on ν,—since $\sum_{k=1}^{\infty}\lambda^k=\lambda/(1-\lambda)$ and $\sum_{k=1}^{\infty}k\lambda^k<+\infty$ if $|\lambda|<1$—the series (*) con-

verges in the sense of (35β). Moreover (since grouping terms never increases sums of absolute values) for the same reasons $\sum_{i=1}^{\infty}\int|d\rho_i^{\nu+1}|$ (which bounds $\sum_{i=1}^{\infty}\sup|\rho_i^{\nu+1}|$) does not exceed the corresponding sum for P^ν by a proportion of more than $\int|d\rho_\nu^\nu|/(1-\int|d\rho_\nu^\nu|)$. And by induction this is at most $5\int|d\rho_\nu^\nu|/4$. Consequently

$$\sum_{i=\nu+1}^{\infty}\int|d\rho_i^{\nu+1}| \leqq \sum_{i=\nu}^{\infty}\int|d\rho_i^{\nu}| - \int|d\rho_\nu^{\nu}| + \frac{1}{5}\left(\frac{5}{4}\int|d\rho_\nu^{\nu}|\right)$$
$$\leqq \sum_{i=\nu}^{\infty}\int|d\rho_i^{\nu}| - \frac{3}{4}\int|d\rho_\nu^{\nu}|$$

and

$$\sum_{i=1}^{\nu}\int|d\rho_i^{\nu+1}| \leqq \sum_{i=1}^{\nu-1}\int|d\rho_i^{\nu}| + \int|d\rho_\nu^{\nu}|.$$

But four-thirds of the first sum, plus the second sum, is non-increasing as $\nu\uparrow\infty$—whence the second sum is always bounded by $\frac{4}{3}\cdot\frac{1}{10}<\frac{1}{5}$, and the first tends to zero.

This proves (36α). Hence (regrouping the terms of (*) through (32α)), by (32α) and (35α), $P^{\nu+1}\sim P^\nu\sim P$. And since by inequalities just proved, $|t((P^{\nu+1})^*)-\sum_{i=1}^{\nu}\gamma_i b_i|$ tends to zero as ν increases, $t(P^*)\equiv t((P^\nu)^*)$ $=\sum_{i=1}^{\infty}\gamma_i\cdot b_i$.

This completes the proof of Theorem 15.

37. **Corollaries of Theorem** 15. Theorem 15 has several immediate corollaries of primary theoretical importance. We shall list some of these now.

COROLLARY 15.1. *One can write $f(x, y)$ as the sum of an infinite series of scalar multiples of brackets of x and y arranged in order of increasing weight; each coefficient can be computed after a finite number of rational operations, and are rational numbers.*

Proof. In Theorem 14, $r\dagger(\lambda)$ is (cf. Remark 2 above) a regular path whose $\rho_i(\lambda)$ are polynomials (of degree at most the length $w(b_i)$ of b_i) with rational numbers as coefficients. These properties are preserved under the rational operations, differentiations, and integrations performed above—any polynomial can be differentiated or integrated by rational operations on its coefficients.

(The reader will find it instructive to compute the terms of degrees two and three.)

Caution. Because of the linear interdependence (due to the identities of Lie-Jacobi) between the brackets of length w, the series of Theorem 15 is not unique; its computation depends on the arrangement of the brackets of each length w.

Corollary 15.2. *The function $x \circ y = f(x, y)$ of composition of any analytical group G under canonical parameters is analytical.*

Proof. By §24, brackets are polynomial functions.

Corollary 15.3. *If the Lie albebra of G is "w-nilpotent" (that is, if all brackets of length w vanish), then $f(x, y)$ is a polynomial of degree at most r.*

Corollary 15.4. *Two analytical groups having topologically isomorphic Lie algebras are locally topologically isomorphic (and so analytically isomorphic).*

Proof. Within some neighborhood of the identity, and under canonical parameters, they have the same function of composition.

Corollary 15.5. *Let L be the Lie algebra of any analytical group G, and let S be any closed subalgebra of L. Then the elements in S near the origin are an analytical subgroup nucleus.*

Proof. They are a subgroup (by Corollary 15.1), satisfy (1), (2), (2′), and are a complete linear subspace of L.

From Corollary 5 and (31α), we get

Corollary 15.6. *The analytical subgroup nuclei of any analytical group G under canonical parameters, are the closed subalgebras of its metric Lie algebra.*

Corollary 15.7. *A locally compact analytical group is a Lie group in the usual sense.*

Proof. Any locally compact Banach space is finite-dimensional by [1], p. 84, and the function of composition is by Corollary 2 analytical under canonical parameters.

Corollary 15.8. *A commutative analytical group nucleus under canonical parameters is a neighborhood of the origin in a Banach space.*

38. **Digression: paths and group-products.** Since to assert $x_1 \circ \cdots \circ x_n = y_1 \circ \cdots \circ y_n$ is to assert

$$P_{x_1} \oplus \cdots \oplus P_{x_n} \sim P_{y_1} \oplus \cdots \oplus P_{y_n},$$

and since every admissible path can be approximated arbitrarily closely by broken lines, one would expect product-equivalences‡ $P \sim Q$ between images of an interval $[0, \Lambda]$ to correspond to algebraic identities between group products. We shall sketch in §38 some crude examples of such correspondences.

The identity $xy = yx(x, y)$ shows that if Q is any broken line, one can replace any two segments of Q by the opposite sides of the parallelogram which

‡ We recall the notation $P \sim Q$ meaning $t(P^*) = t(Q^*)$.

they determine, without altering $t(Q^*)$, provided a small deviation (x, y) is inserted.

The graphical principle (essential in the classical proofs of Green's and Stokes' Theorems) that any path-deformation can be split up into elementary deformations across parallelograms, is analogous to the algebraic principle that any permutation of terms in a sequence is the product of transpositions.

The derivation in §34, given a path Q, of paths $P \sim Q$ by choosing $v^*(0) = v^*(\Lambda) = 0$ and setting

$$dp = v^{*-1}(\lambda) \circ dq \circ v^*(\lambda) + dv$$

corresponds to taking a product $x_1 \circ \cdots \circ x_n$ and a second product $u_1 \circ \cdots \circ u_n = e$, defining $v_k^* = u_1 \circ \cdots \circ u_k$ and proving by induction $\prod_{i=1}^{k} [(v_{i-1}^{*-1} \circ x_i \circ v_{i-1}^*) \circ u_i] = x_1 \circ \cdots \circ x_k \circ v_k^*$, and thus concluding that

$$\prod_{k=1}^{n} x_k = \prod_{k=1}^{n} [(v_{k-1}^{*-1} \circ x_k \circ v_{k-1}^{*}) \circ u_k]. \tag{38α}$$

39. **Digression: the Rearrangeability Principle.** In correlating the argument of §§34–36 with formal identities on group products, let us begin by recalling a recent result of P. Hall ([7], Theorem 3.1), namely

$$(xy)^n \equiv x^n y^n z_1^{\phi_1(n)} \cdots z_t^{\phi_t(n)} \pmod{H_w}, \tag{39α}$$

where the z_k are complex commutators in x and y of lengths $<w$ arranged in order of increasing length, the exponents ϕ_k are polynomials of degree $w(z_k)$, and H_w is the normal subgroup whose elements are the products of commutators of lengths $\geqq w$.

That there exist (not necessarily polynomial) *functions* $\phi_k(n)$ such that (39α) is satisfied, is very easy to show. For since $uv = vu(u, v)$, one can transpose any two adjacent terms in any product involving x, y, and their commutators, by inserting commutators of lengths greater than the length of either transposed term. Hence one can first shift all the occurrences of x in such a product to the extreme left, then all the occurrences of y to positions just to the right of these, and similarly with $z_1, \cdots, z_t$.

This method, combined with the rule that any permutation can be accomplished by successive transpositions, obviously yields a general

Rearrangeability Principle. If one is given *any* product ψ involving elements $x_1, \cdots, x_n$ and their commutators, *any* integer w, and *any* ordering ρ of the x_k and their commutators of weights $<w$, then ψ is congruent modulo commutators of weights $\geqq w$ to a product of powers of the x_k and their commutators arranged in the sequence ρ.

More than this, one can *distribute* their occurrences according to any preassigned distribution function.

These principles are the key to the algebraic situation. Using them, one can show for instance that if $m=n^w$, then

$$x^my^m \equiv \prod_{k=1}^{n} v_k \pmod{H_w}, \tag{39β}$$

where each v_k is of the form $x^{m/n}y^{m/n}z_1^{\zeta_1(m;k)}\cdots z_t^{\zeta_t(m;k)}$ and $|\zeta_h(m;i)-\zeta_h(m;j)| \leqq 1$, which means that the v_k are all nearly equal.

Proof. Write $x^my^m=(x^{m/n})(x^{m/n})\cdots(x^{m/n})(x^{m/n}y^m)$. Then transpose the occurrences of y (inserting commutators, of course) until you obtain the identity

$$x^my^m = (x^{m/n}y^{m/n}u_1)(x^{m/n}y^{m/n}u_2)\cdots(x^{m/n}y^{m/n}u_3),$$

where the u_k are congruent (mod H_w) to products of the commutators z_k of lengths $<w$. Proceed by induction, dividing the occurrences of each z_h $(h=1, \cdots, t)$ into n nearly equal lots, and you will get (39β).

Now suppose x and y are elements of a continuous group under canonical parameters. Write $x^m=\bar{x}$ and $y^m=\bar{y}$; since the v_k are nearly equal, if we know by (27β) that the elements of H_w are relatively small, we see that $|f(\bar{x}, \bar{y})-nv_1|$ is small, where for n large v_1 is nearly determined by x, y, and the commutation function $[x, y]$.

40. **Digression (cont.): analyticity and other remarks.** We do not have to go far beyond the same principles to see from an algebraic standpoint even why an SCH-series exists, in the way that it does.

To see this, observe that for fixed $x^m=\bar{x}$, $y^m=\bar{y}$ and very large n, since x and y are correspondingly small, (1) products are nearly sums, and (2) commutators are nearly equal to the corresponding brackets. Hence if b_h denotes the bracket in $\bar{x}$ and $\bar{y}$ corresponding to the commutator in x and y denoted by z_h, and $\lambda_h=n\zeta_h(m;1)/n^{w(z_h)}$, then the smallness of $|f(\bar{x}, \bar{y})-nv_1|$ implies the smallness of $|f(\bar{x}, \bar{y})-\{\bar{x}+\bar{y}+\sum_{h=1}^{t}\lambda_h b_h\}|$. This gives one the first $(t+2)$ terms of an SCH-series, approximately.

Actually, the λ_h are polynomials whose dominating terms are independent of m, although the reasons for this are number-theoretical and not at all trivial, and the calculation of the dominating terms is not even impossibly laborious.

Similar reasoning yields an algebraic paraphrase of Theorem 15. Take any path X: $x(\lambda)=\sum_{i=1}^{r}\rho_i(\lambda)\cdot x_i$. Divide X into $m=n^w$ equal parts, set

$\mu_i^k = \rho_i(k,n) - \rho_i([k-1]/n)$ and $x = y^m$, and obtain through the Rearrangeability Principle (as with (39β)), an identity‡

$$(40\alpha) \qquad \prod_{k=1}^{n}\left(\prod_{i=1}^{r} y_i^{[m\mu_i^k]}\right) \equiv \prod_{k=1}^{n} v_k \pmod{H_w},$$

where the v_k are products $\prod_{h=1}^{t} z_h^{\xi_h(m;k)}$ of nearly equal powers of commutators z_h in the y_i. But replacing each commutator z_h of length w_h in the y_i by $(m^{-w_h})\cdot b_h$, where b_h denotes the bracket in the x_i corresponding to z_h—the substitution is nearly one of equals for equals—and setting $\gamma_h(m) = nm^{-w_h}\zeta_h(m;1)$, (40$\alpha$) becomes

$$(40\alpha) \qquad t(X^*) \text{ is nearly } \sum_{h=1}^{t} \gamma_h(m)\cdot b_h,$$

the calculation of $\gamma_h(m)$ being the same for all groups.

41. **Every metric Lie algebra belongs to a group.** We can now prove by considerations of convergence, that

THEOREM 16. *Every metric Lie algebra L is the Lie algebra of an analytical group nucleus.*

Proof. Define group products $x \circ y = f(x, y)$ in L through the SCH-series. There are three points to establish: the convergence of the series, the validity of the inequalities (2′)–(2″), and the associative law $f(f(x, y), z) = f(x, f(y, z))$.

By Remark 2 of §30, we can assume $|[x, y]| \leq |x|\cdot|y|$. Then by the proof of Theorem 15 (cf. the remark after (36α)), if we substitute for each bracket in the SCH-series, the product of the absolute values of its entries, and if these are $<1/10$, then the sum of the absolute values of the resulting series is bounded by $2(|x|+|y|)$. The *convergence* of $f(x, y)$ provided $|x|+|y| < 1/10$ is a weak corollary of this.

Again, expanding $[f(x, a) - f(x, b)] - (a - b)$ in SCH-series, we have by Theorem 14 after cancellation and pairing off of corresponding terms, scalar multiples of differences such as

$$\begin{aligned}\Phi &= [x, [a, x], a] - [x, [b, x], b]\\ &= [x, [a - b, x], a] + [x, [b, x], a - b]\end{aligned}$$

whose magnitude is bounded by $|x|\cdot|a-b|$ times the number n_i of entries in the bracket, times what we would get if we replaced every bracket by the product of the absolute values of all but one of its entries. But the sum

‡ $[m\mu_i^k]$ denotes conventionally the integral part of $m\mu_i^k$.

of the products of these last two factors still converges absolutely provided $|a|+|b|+|x|<1/20$, so that $n_i(1/20)^{n_i-1}<10(1/10)^{n_i}$. Hence

$$|(xa - xb)| - |(a - b)| \leqq K\cdot |x| \cdot |a - b|$$

within this region. This implies (2′); (2″) follows by symmetry. It remains to prove the associative law.

Here we do the obvious thing: substitute the *SCH*-series for u in the *SCH*-series for $f(u, z)$, and likewise the *SCH*-series for $f(y, z)$ for v in the *SCH*-series for $f(x, v)$, and expand in both cases by the distributive law. We will get two series of monomial brackets in x, y, z, with possible repetitions. If they are absolutely convergent, then by the continuity implied in (2′)–(2″) they will converge to $f((x, y), z)$ and $f(x, f(y, z))$ respectively. We shall next prove that they are absolutely convergent.

If $|x|+|y|+|z|<1/80$, and we replace each bracket in the series for $f(f(x, y), z)$ by the product of the absolute values of its entries, then the sum of the absolute values of what we get is by the distributive law (on scalars) what we *would* get if we replaced brackets by products in the *SCH*-series for $f(u, z)$, replaced z by $|z|$, and u by the sum of the absolute values of the terms in the *SCH*-series for $f(x, y)$. And since both of these are $<1/40$, the series for $f(f(x, y), z)$ is absolutely convergent. The absolute convergence of the series for $f(x, f(y, z))$ follows by symmetry.

Hence to prove that $f(f(x, y), z)=f(x, f(y, z))$ we need only show that irrespective of n, the sum of the terms of length $\leqq n$ is the same for the two series. The demonstration of this essentially algebraic fact completes the proof.

Demonstration. Form the multiplicative group of all non-commutative polynomials $I+U=I+\lambda_1X+\lambda_2Y+\lambda_3Z+\cdots$ in X, Y, Z, ignoring terms of degree $>n$. This is a (4^n-1)-parameter Lie group, in which $(I+U)^{-1}=I-U+U^2-\cdots+(-1)^nU^n$. Since the group is analytical, the functions $f(f(X, Y), Z)$ and $f(X, f(Y, Z))$ are identically equal near $X=Y=Z=0$, and hence formally equal. Moreover as in all linear groups, $[U, V]=VU-UV$. But by Theorem 3 of the author's *Representability of Lie algebras and Lie groups by matrices*, Annals of Mathematics, vol. 38 (1937), pp. 526–532, any identity between alternants $VU-UV$ follows formally from the identities of Lie-Jacobi. Hence the equality between the sums of the terms of degree $\leqq n$ in the two series follows formally from the identities of Lie-Jacobi (which we assumed at the beginning).

BIBLIOGRAPHY

1. S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.
2. G. Birkhoff, *Integration of functions with values in a Banach space*, these Transactions, vol. 38 (1935), pp. 357–378.
3. J. E. Campbell, *On a law of combination of operators*, Proceedings of the London Mathematical Society, vol. 28 (1897), pp. 381–390.
4. E. Cartan, *Notice sur les Travaux Scientifiques de M. Élie Cartan*, Chap. VIII, Paris, 1931.
5. E. Cartan, *Les groupes continus et l'analysis situs*, Mémorials des Sciences Mathématique, Fasc. 42, Paris, 1930.
6. J. Delsarte, *Les groupes de transformations linéaires dans l'espace de Hilbert*, Mémorials des Sciences Mathématiques, Fasc. 57, Paris, 1932.
7. P. Hall, *A contribution to the theory of groups of prime-power orders*, Proceedings of the London Mathematical Society, vol. 36 (1933), pp. 29–95.
8. F. Hausdorff, *Die symbolische Exponentialformel in der Gruppentheorie*, Leipzig Berichte, vol. 58 (1906), pp. 19–48.
9. S. Lie, *Transformationsgruppen*, Leipzig, 1888.
10. S. Lie, *Unendlich continuirliche Gruppen*, Abhandlungen, Sachsische Gesellschaft der Wissenschaften, vol. 21 (1895), pp. 43–150.
11. W. Mayer and T. Y. Thomas, *Foundations of the theory of continuous groups*, Annals of Mathematics, vol. 36 (1935), pp. 770–782.
12. J. von Neumann, *Gruppen linearen Transformationen*, Mathematische Zeitschrift, vol. 30 (1929), pp. 3–42.
13. S. Saks, *Théorie de l'Intégrale*, Warsaw, 1933.
14. L. Schlesinger, *Neue Grundlagen für eine Infinitesimalkalkul der Matrizen*, Mathematische Zeitschrift, vol. 33 (1931), pp. 33–61.
15. O. Schreier, *Abstrakte kontinuierliche Gruppen*, Hamburger Abhandlungen, vol. 4 (1926), pp. 15–32.
16. F. Schur, *Neue Begründung der Theorie der endlichen Transformations gruppen*, Mathematische Annalen, vol. 35 (1889), pp. 161–197.
17. V. Volterra, *Sulle equazione differenziali lineari*, Rendiconti dei Lincei, vol. 3 (1887), pp. 391–396.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

ON AN APPROXIMATE FUNCTIONAL EQUATION OF PALEY*

BY

W. C. RANDELS†

The purpose of this paper is to present and extend a method which has been found useful in the construction of "gegenbeispiels." The methods of proof have all been used before by various writers, but the conditions used here are more general than those considered by the other writers.

If we have a sequence of positive numbers $\{a_n\}$ such that

$$\sum_{n=1}^{\infty} a_n = \infty$$

and such that the function,

$$f(z) = \sum_{n=1}^{\infty} a_n z^n,$$

is analytic inside the unit circle, we know that $f(\rho e^{i\theta})$ will tend to infinity as $\rho \to 1$, at least for $\theta = 0$. If we introduce a sequence of factors $\{e^{ib_n}\}$ (b_n real), the function

$$f(z) = \sum_{n=1}^{\infty} a_n e^{ib_n} z^n$$

will tend to have its singularities spread over the unit circle and the order of $\max_{0 \leq \theta < 2\pi} f(\rho e^{i\theta})$, as $\rho \to 1$, will be decreased. This suggests the problem of determining a sequence $\{b_n\}$ corresponding to the sequence $\{a_n\}$ such that the order of $f(\rho e^{i\theta})$ as $\rho \to 1$ is the same for all values of θ.

The first result of this nature was given by Hardy and Littlewood [2].‡ They considered the case of $a_n = n^{\beta - 1/2}$ and they proved that

$$f_\beta(z) = \frac{Ht^{-\beta}}{\log a} e^{-\pi i/4} \sum_{n=1}^{\infty} n^{\beta - 1/2} \exp\,[i\alpha n \log n] z^n = F(\sigma) + \phi(\sigma), \S$$

where

---

* Presented to the Society, February 23, 1935; received by the editors February 11, 1937.

† This work was done while the author was a Sterling Research Fellow at Yale University. The problem was suggested by Professor Hille.

‡ The numbers in brackets refer to the references at the end of the paper.

§ To make the printing easier we shall sometimes write exp $[x]$ instead of e^x.

$$\alpha = \frac{2\pi}{\log a}, \quad H = \frac{(2\pi)^\beta}{(\log a)^{\beta+1/2}}, \quad z = \rho e^{i\theta}, \quad \rho = e^{\alpha\sigma/t}, \quad \theta = \alpha \log\left(\frac{\alpha}{et}\right),$$

and

$$F(\sigma) = \sum_{\nu=1}^{\infty} a^{\beta\nu} e^{-(\sigma+it)a^\nu},$$

$$\phi(\sigma) = \begin{cases} A + o(1), \text{ as } \sigma \to 0, & \text{if } \beta < \frac{1}{2}, \\ O\left(\log \frac{1}{\sigma}\right), & \text{if } \beta = \frac{1}{2}, \\ O(\sigma^{-\beta+1/2}), & \text{if } \beta > \frac{1}{2}. \end{cases}$$

The function $F(\sigma)$ is similar to the Weierstrass non-differentiable function, and extensive work by Hardy [1] on that function makes it possible to show that $f_\beta(z)$ is continuous for $|z| = 1$ and $\beta < 0$, and

$$(1) \qquad f_\beta(z) = \begin{cases} O\{(1 - |z|)^{-\beta}\}, & \beta > 0, \\ O\left\{\log\left(\frac{1}{1 - |z|}\right)\right\}, & \beta = 0. \end{cases}$$

Moreover they showed that*

$$(2) \qquad f_\beta(z) = \Omega\{(1 - |z|)^{-\beta}\} \text{ as } \rho \to 1, \qquad \beta > 0.$$

As a consequence of (1) and a theorem of Hardy and Littlewood [4, Theorem 4] it follows that

$$(3) \quad \omega(f_\beta, h) = \underset{|\theta_1 - \theta_2| \le h}{\text{l.u.b.}} \; |f_\beta(e^{i\theta_1}) - f_\beta(e^{i\theta_2})| = \begin{cases} O(h^{-\beta}), & -1 < \beta < 0, \\ O\left(h \log \frac{1}{h}\right), & \beta = -1. \end{cases}$$

Hille [5] has given a proof of (1), but by his method one cannot prove (2). Hille, and Ingham [6], considered the similar functions

$$f_{\beta\gamma}(z) = \sum_{n=2}^{\infty} n^{-1/2}(\log n)^{-\gamma} e^{in(\log n)^\beta} z^n, \qquad \beta < 1, \qquad \gamma > \tfrac{1}{2}(\beta + 1),$$

and showed that these functions are continuous for $|z| = 1$.

It is natural to try to prove similar theorems for functions of the type

$$f(z) = \sum_{n=1}^{\infty} b(n) e^{i\Delta(n)} z^n.$$

It is convenient to assume that the functions $b(x)$ and $\Delta(x)$ are defined for

* By the notation $f(x) = \Omega\{g(x)\}$ we mean that $f(x) \neq o\{g(x)\}$.

all values of x on the interval $(1, \infty)$ and we shall do so from now on. This problem was first handled by Paley [7] who showed that, if $n_\nu(\theta)$ is defined by the relation

$$\Delta'\{n_\nu(\theta)\} = 2\pi\nu - \theta,$$

then

$$\begin{aligned} f(z) &= \sum_{n=1}^{\infty} b(n)e^{i\Delta(n)}z^n \\ &= \pi^{1/2}e^{\pi i/4}\sum_{\nu=1}^{\infty} \frac{b\{n_\nu(\theta)\}}{(\Delta''\{n_\nu(\theta)\})^{1/2}}\,|z|^{n_\nu(\theta)} \exp\left[i\{\Delta[n_\nu(\theta)] + \theta n_\nu(\theta)\right. \\ &\qquad\left. - 2\pi\nu n_\nu(\theta)\}\right] + R(z), \end{aligned} \tag{4}$$

where $R(z)$ is continuous for $|z| = 1$. Paley's results however were not as general as ours for he used a condition of the form

$$b(n) = O(x^{\epsilon-1/2}), \qquad \epsilon < 1/10.$$

Wilton [8] working on the same series when $|z| = 1$ proved the equation

$$\begin{aligned} f_m(e^{i\theta}) &= \sum_{n=1}^{m} b(n)e^{i\Delta(n)}e^{in\theta} \\ &= \pi^{1/2}e^{\pi i/4}\sum_{\nu=1}^{\mu} \frac{b\{n_\nu(\theta)\}}{(\Delta''\{n_\nu(\theta)\})^{1/2}} \exp\left[i\{\Delta[n_\nu(\theta)] + \theta n_\nu(\theta) - 2\pi n_\nu(\theta)\}\right] \\ &\qquad + O(1) + o\left(\int_1^{\mu} \frac{b(x)[\Delta''(x)]^{1/2}}{\Delta'(x)}\,dx\right), \end{aligned} \tag{5}$$

where $n_\mu(\theta) \leqq m < n_{\mu+1}(\theta)$. Wilton mentions that his equation is in general a special case of one of van der Corput. We notice that neither the method of Paley nor that of Wilton will prove (1) or (2).

The purpose of this paper is to give an extension of Paley's methods so as to include (1) as a special case. In §1 we prove a functional equation similar to (4) except that $R(z)$ is not necessarily continuous. The functional equation of §1 could possibly be gotten from Wilton's (5) by replacing $b(x)$ by $\rho^x b(x)$ but it would involve at least as many complications as the proof given here.

In §2, it is shown by a method of Hardy [1] that

$$\begin{aligned} G(z) &= \pi^{1/2}e^{\pi i/4}\sum_{\nu=1}^{\infty} \rho^{n_\nu} \frac{b(n_\nu)}{[\Delta''(n_\nu)]^{1/2}} \exp\left[i\{\Delta(n_\nu) + \theta n_\nu - 2\pi\nu n_\nu\}\right] \\ &= O\left\{M\left(\frac{1}{1-\rho}\right)\right\}, \end{aligned}$$

where, if $n_\mu(0) \leqq x < n_{\mu+1}(0)$,

$$M(x) = \sum_{\nu=1}^{\mu} \max_{n_{\nu-1}(0) \leqq y \leqq n_\nu(0)} \frac{b(y)}{[\Delta''(y)]^{1/2}} + \max_{n_\mu(0) \leqq y \leqq x} \frac{b(y)}{[\Delta''(y)]^{1/2}}.$$

We also show that

$$G(z) = \Omega\left\{M\left(\frac{1}{1-\rho}\right)\right\},$$

if

$$\text{(D)} \qquad M(x) = O\left\{\frac{b(x)}{[\Delta''(x)]^{1/2}}\right\}.$$

The remainder function $R(z)$ is studied in §3 and it is shown that if $M(x)$ is bounded, $R(z)$ is continuous for $|z| = 1$ and otherwise

$$R(z) = o\left\{M\left(\frac{1}{1-\rho}\right)\right\}.$$

In §4 we use a method of Hardy and Littlewood [4] to show that if $f(z)$ is analytic for $|z| < 1$ and continuous for $|z| = 1$, and

$$f'(z) = O\{g(|z|)\},$$

where $g(x) \uparrow \infty$ as $x \to 1$, then

$$\omega(f, h) = \underset{|\theta_1 - \theta_2| \leqq h}{\text{l.u.b.}} |f(e^{i\theta_1}) - f(e^{i\theta_2})| = O\left\{\int_{1-h}^{1} g(x)dx\right\},$$

provided the integral exists.

Wilton [8] has given some applications of his equation so that in discussing applications we confine ourselves to cases which have not been covered by Wilton. It is pointed out that (1), (2), and (3) follow from this theory and also certain generalizations of them are possible. Ingham's results also follow as pointed out by Wilton, and furthermore

$$\omega(f_{\beta\gamma}, h) = O\left\{\left(\log \frac{1}{h}\right)^{-\gamma-\beta/2+3/2}\right\}, \quad \text{if} \quad \gamma + \frac{\beta}{2} > \frac{3}{2},$$

where $f_{\beta\gamma}(z)$ is defined by

$$f_{\beta\gamma}(z) = \sum_{n=2}^{\infty} n^{-1/2}(\log n)^{-\gamma} \exp\left[i\int_1^n (\log x)^\beta dx\right].$$

It is finally shown that, if $b(x)$ satisfies conditions (A), (B), and (C) of §1, and

$$\text{(E)} \qquad \int_1^x b^2(x)dx = O\{x^{1+\epsilon}b^2(x)\}, \qquad \epsilon < 1/30,$$

and

$$\int_1^\infty b^2(x)dx = \infty,$$

then, if $\Delta(x)$ is defined by

$$\Delta(x) = \int_2^x \log\left[\int_1^y b^2(z)dz\right]dy,$$

the functions $\Delta(x)$ and $b(x)$ will satisfy condition (D) of §2 which implies that the singularities of the function

$$f(z) = \sum_{n=2}^{\infty} b(n)e^{i\Delta(n)}z^n$$

are distributed uniformly over the unit circle, and the order of $f(\rho e^{i\theta})$ as $\rho\to 1$ is the same for all values of θ. This answers a question proposed in the opening paragraph.

1. **The functional equation.** The method used in this section is essentially that of Paley and the notation is chosen to agree with his. If the functions $b(x)$ and $d(x)$ are given we define

$$\begin{aligned} \Delta(x) &= \int_1^x d(y)dy, \\ z &= \rho e^{i\theta}, \\ d\{n_\nu(\theta)\} &= 2\pi\nu - \theta, \qquad 0 \leqq \theta < 2\pi, \qquad \nu = 1, 2, \cdots, \\ F(\nu;\theta) &= \Delta\{n_\nu(\theta)\} + \theta n_\nu(\theta) - 2\pi\nu n_\nu(\theta). \end{aligned}$$

We impose the following conditions on $b(x)$ and $d(x)$:*

(A) The function $d(x)$ has two derivatives, continuous on the open interval $(1, \infty)$ and

$$\begin{aligned} d(x) &\uparrow \infty \text{ as } x \to \infty, \\ d'(x) &= O(x^{-1}), \\ 1/d'(x) &= O(x^{1+\epsilon}), \qquad 0 < \epsilon < 1/30, \\ d''(x) &= O(x^{\eta-2}), \qquad \eta < 1/10 - \epsilon. \end{aligned}$$

We notice that the first condition on $d(x)$ implies that $n_\nu(\theta) \uparrow \infty$ as $\nu \to \infty$.

* It would obviously be sufficient to assume that these conditions are satisfied only for x greater than a certain x_0; but to simplify the work we carry through the proof for $x_0 = 1$.

(B) The function $b(x)$ is monotone and has one derivative continuous on $(1, \infty)$ and

$$b'(x) = O\{b(x)x^{\delta-1}\}, \qquad \delta \leqq \epsilon.$$

The function $d'(x)$ and $b(x)$ satisfy the condition:

(C) There exists a constant $k>1$, such that for each function

$$f(x') < kf(x), \qquad kf(x') > f(x), \quad \text{if} \quad 1 \leqq x \leqq x' \leqq ex.$$

If we had restricted ourselves to functions of the logarithmic-exponential type, condition (C) would mean that the functions could not be like e^x or e^{-x}. We notice some consequences of (C). If $f(x)$ satisfies (C), then

$$\frac{1}{f(x')} < \frac{k}{f(x)}, \qquad \frac{k}{f(x')} > \frac{1}{f(x)}, \qquad 1 \leqq x \leqq x' \leqq ex.$$

If two functions $f_1(x)$ and $f_2(x)$ satisfy (C) with constants k_1 and k_2 respectively, then for $x \leqq x' \leqq ex$,

$$f_1(x')f_2(x') < k_1k_2f_1(x)f_2(x), \qquad k_1k_2f_1(x')f_2(x') > f_1(x)f_2(x).$$

If $f(x)$ satisfies (C), $f(x)>0$ then for $x \leqq x' \leqq ex$,

$$k\int_{x'}^{A} f(y)dy > \int_{x}^{A+x-x'} f(y)dy, \qquad k\int_{x'}^{\infty} f(y)dy > \int_{x}^{\infty} f(y)dy,$$

and

$$\int_{e}^{x'} f(y)dy < ek\int_{1}^{x} f(y)dy,$$

$$\int_{1}^{x'} f(y)dy < ek\int_{1}^{x} f(y)dy + \int_{1}^{e} f(y)dy.$$

This means that, if the functions $f_1(x)$ and $f_2(x)$ satisfy (C) and are positive, then the functions

$$\frac{1}{f_1(x)}, \qquad f_1(x)f_2(x), \qquad \int_{x}^{\infty} f_1(y)dy, \qquad \int_{1}^{x} f_1(y)dy$$

also satisfy (C) for $x>a>1$ for some a. If $f(x)>0$ satisfies (C), then for $e^n x \leqq x' < e^{n+1}x$,

$$\frac{f(x')}{f(x)} = \frac{f(ex)}{f(x)} \cdots \frac{f(x')}{f(e^n x)} < k^{n+1} < k \cdot k^{\log(x'/x)} = k\left(\frac{x'}{x}\right)^{\log k}.$$

If $f(x)>0$ satisfies (C), then

$$\sum_{n=1}^{\infty} f(n) < k \sum_{n=1}^{\infty} \int_{n}^{n+1} f(x)dx = k \int_{1}^{\infty} f(x)dx$$

and

$$\int_{1}^{\infty} f(x)dx = \sum_{n=1}^{\infty} \int_{n}^{n+1} f(x)dx < k \sum_{n=1}^{\infty} f(n),$$

so that the series and integrals are equiconvergent.

We wish to obtain an approximate functional equation for the function,

$$f(z) = \sum_{n=1}^{\infty} b(n) \exp [i\Delta(n)]z^n.$$

We first consider the group of terms*

$$\sum_{r-s}^{r+s} b(n) \exp [i\Delta(n)]z^n = \sum_{r-s}^{r+s} c(n) \exp [i\Delta(n) + in\theta],$$

where $r = n_\nu$, $s = n_\nu^{3/5}$, and where $c(x) = b(x)\rho^x$. We have

$$\sum_{r-s}^{r+s} c(n) \exp [i\Delta(n) + in\theta] - c(n_\nu) \sum_{r-s}^{r+s} \exp [i\Delta(n) + in\theta]$$

$$(1.1) \qquad \begin{aligned} &= O\left\{ \sum_{-s}^{s} | c(n + n_\nu) - c(n) | \right\} \\ &= O\left\{ \max_{|x - n_\nu| \leqq s} \left| \frac{d}{dx} c(x) \right| \sum_{0}^{s} n \right\} \\ &= O\{(1 - \rho)b(n_\nu)\rho^{n_\nu/2} n_\nu^{6/5} + n_\nu^{1/5+\delta} b(n_\nu)\rho^{n_\nu/2}\}, \end{aligned}$$

since by (C) and (B),

$$c'(x) = b(x)\rho^x \log \rho + b'(x)\rho^x = O\{b(n_\nu)\rho^{n_\nu/2}(1 - \rho) + n_\nu^{\delta-1} b(n_\nu)\rho^{n_\nu/2}\},$$
$$\tfrac{1}{2}n_\nu \leqq x \leqq \tfrac{3}{2}n_\nu.$$

By (A)

$$\sum_{r-s}^{r+s} \exp [i\Delta(n) + in\theta] - \int_{r-s}^{r+s} \exp [i\Delta(x) + ix\theta - 2\pi i\nu x]dx$$

$$(1.2) \qquad \begin{aligned} &= O\left\{ n_\nu^{3/5} \max_{|x - n_\nu| \leqq s} \left| \frac{d}{dx} [i\Delta(x) + ix\theta - 2\pi i\nu x] \right| \right\} \\ &= O\left\{ n_\nu^{3/5} \max_{|x - n_\nu| \leqq s} \left| \int_{n_\nu}^{x} d'(y)dy \right| \right\} \end{aligned}$$

---

* By the symbol $\sum_a^b$ we shall understand the sum over all values of n such that $a \leqq n < b$.

$$= O\left\{n_\nu^{6/5} \max_{|x-n_\nu|\leq s} d'(x)\right\} = O\{n_\nu^{1/5}\}.$$

We have, again using (A),

$$\Delta(x) + x\theta - 2\pi\nu x = F(\nu;\theta) + (x - n_\nu)^2 d'(n_\nu) + O\left\{\frac{|x - n_\nu|^3}{n_\nu^{2-\eta}}\right\},$$

so that

$$\int_{r-s}^{r+s} \exp[i\Delta(x) + ix\theta - 2\pi i\nu x]dx - \exp[iF(\nu;\theta)]\int_{r-s}^{r+s} \exp[i(x - n_\nu)^2 d'(n_\nu)]dx$$

$$= O\left\{\int_{-s}^{s}\left|1 - \exp\left[iO\left(\frac{x^3}{n_\nu^{2-\eta}}\right)\right]\right| dx\right\}$$

$$= O\left\{\frac{1}{n_\nu^{2-\eta}}\int_{-s}^{s}|x^3|\,dx\right\} = O\{n_\nu^{2/5+\eta}\}.$$

By a simple change of variable

$$\int_{r-s}^{r+s} \exp[i(x - n_\nu)^2 d'(n_\nu)]dx = 2[d'(n_\nu)]^{-1/2}\int_0^t e^{ix^2}dx,$$

where $t = n_\nu^{3/5}[d'(n_\nu)]^{-1/2}$, and we know that

$$\int_0^\infty e^{ix^2}dx = \tfrac{1}{2}e^{\pi i/4}\pi^{1/2}.$$

By considering the graph of the function $\cos x^2$ we see that $\int_a^\infty \cos x^2\,dx$ can be represented as an alternating series for which the absolute value of the terms is steadily decreasing. Therefore

$$\int_t^\infty \cos x^2\,dx = O\{[d'(n_\nu)]^{1/2} n_\nu^{-3/5}\}.$$

By applying similar considerations to $\sin x^2$ we see that

$$\int_t^\infty e^{ix^2}dx = O\{[d'(n_\nu)]^{1/2} n_\nu^{-3/5}\}.$$

Hence

$$(1.3)\quad \int_{r-s}^{r+s} \exp[i\Delta(x) + ix\theta - 2\pi i\nu x]dx - \pi^{1/2}e^{\pi i/4}[d'(n_\nu)]^{-1/2} \exp[iF(\nu;\theta)]$$
$$= O\{n_\nu^{2/5+\eta} + n_\nu^{-3/5}\} = O\{n_\nu^{2/5+\eta}\}.$$

This implies that

$$\sum_{r-s}^{r+s} c(n) \exp\,[i\Delta(n) + in\theta] = \pi^{1/2}e^{\pi i/4}[d'(n_\nu)]^{-1/2} \exp\,[iF(\nu;\theta)]b(n_\nu)\rho^{n_\nu} \tag{1.4}$$
$$+ O\{(1-\rho)b(n_\nu)n_\nu^{6/5}\rho^{n_\nu/2} + b(n_\nu)n_\nu^{1/5+\delta}\rho^{n_\nu/2} + b(n_\nu)n_\nu^{2/5+\eta}\rho^{n_\nu}\}.$$

We set $a(x)=\Delta(x)+\theta x$, $\gamma(x)=a(x+1)-a(x)$. We have then

$$\sum_{r+s}^{r_1-s_1} b(n)e^{ia(n)}\rho^n = \sum_{r+s}^{r_1-s_1} \frac{c(n)}{e^{i\gamma(n)}-1}\{e^{i\gamma(n)}-1\}e^{ia(n)},$$

where $r_1=n_{\nu+1}$, $s_1=n_{\nu+1}^{3/5}$, and if we define

$$H(x) = \sum_{1}^{x} \{e^{i\gamma(n)}-1\}e^{ia(n)} = e^{ia([x+1])} - e^{ia(1)} = O(1),$$

this gives us

$$\sum_{r+s}^{r_1-s_1} b(n)e^{ia(n)}\rho^n = \int_{r+s}^{r_1-s_1} \frac{c(x)}{e^{i\gamma(x)}-1}\,dH(x).$$

Integrating by parts we have

$$\int_{r+s}^{r_1-s_1} \frac{c(x)}{e^{i\gamma(x)}-1}\,dH(x) = \frac{c(x)H(x)}{e^{i\gamma(x)}-1}\Bigg|_{r+s}^{r_1-s_1}$$
$$- \int_{r+s}^{r_1-s_1} H(x)\frac{d}{dx}\left\{\frac{c(x)}{e^{i\gamma(x)}-1}\right\}dx.$$

By the definition of $\gamma(x)$ and property (A),

$$\gamma(x) = d(x) + O\{d'(x)\} + \theta, \qquad \gamma'(x) = d'(x) + O\{d''(x)\}$$

and

$$\gamma(n_\nu) - 2\pi\nu + \theta = O\{d'(n_\nu)\}.$$

Consequently by (C) there exists a constant $M>0$ so that for ν sufficiently large

$$\gamma(n_\nu + n_\nu^{3/5}) - 2\pi\nu + \theta = \gamma(n_\nu) - 2\pi\nu + \theta + \int_{n_\nu}^{r+s}\gamma'(x)dx$$
$$> n_\nu^{3/5}\min_{0\leq x\leq s}[d'(n_\nu+x)] - O\{d'(n_\nu) + n_\nu^{3/5}d''(n_\nu)\}$$
$$> Mn_\nu^{3/5}(n_\nu+n_\nu^{3/5})^{-1-\epsilon},$$

and by (A) for $x>n_\nu+n_\nu^{3/5}$, there exists an M' such that

$$\gamma(x) - 2\pi\nu + \theta > Mn_\nu^{3/5}(n_\nu + n_\nu^{3/5})^{-1-\epsilon} + (x - n_\nu - n_\nu^{3/5}) \min_{r+s \leqq y \leqq x} [d'(y)]$$
$$> M'(x - n_\nu)x^{-1-\epsilon}.$$

If we define N_ν by the relation

$$d(N_\nu) = (2\nu + 1)\pi - \theta,$$

then since $\gamma'(x) = d(x+1) - d(x) > 0$, we have for $x < N_\nu$,

$$|\gamma(x) - 2(\nu + 1)\pi + \theta| \geqq \pi.$$

We know however that, for $2\pi\nu < y \leqq (2\nu+1)\pi$,

$$(e^{iy} - 1)^{-1} = O\{(y - 2\pi\nu)^{-1}\},$$

and therefore

$$(1.5) \quad (e^{i\gamma(x)} - 1)^{-1} = O\{x^{1+\epsilon}(x - n_\nu)^{-1}\}, \qquad n_\nu + n_\nu^{3/5} \leqq x \leqq N_\nu.$$

By a similar reasoning

$$(1.6) \quad (e^{i\gamma(x)} - 1)^{-1} = O\{n_{\nu+1}^{1+\epsilon}(n_{\nu+1} - x)^{-1}\}, \qquad N_\nu \leqq x \leqq n_{\nu+1} - n_{\nu+1}^{3/5},$$

and

$$(1.7) \quad (e^{i\gamma(x)} - 1)^{-1} = O\left\{\left(\frac{n_{\nu+1}}{2^j}\right)^{1+\epsilon}\left(\frac{n_{\nu+1}}{2^j} - x\right)^{-1}\right\}, \qquad N_\nu \leqq x \leqq \frac{n_{\nu+1}}{2^j},$$
$$j = 1, 2, \cdots.$$

Therefore, by (C),

$$(1.8) \qquad \frac{c(x)H(x)}{e^{i\gamma(x)} - 1}\Bigg|_{r+s}^{r_1-s_1} = O\{c(n_\nu)n_\nu^{2/5+\epsilon} + c(\tfrac{1}{2}n_{\nu+1})n_{\nu+1}^{2/5+\epsilon}\}.$$

We also have*

$$\left|\int_{r+s}^{r_1-s_1} H(x)\frac{d}{dx}\left\{\frac{c(x)}{e^{i\gamma(x)} - 1}\right\}dx\right|$$
$$\leqq 2\left[\int_{r+s}^{2n_\nu} + \int_{2n_\nu}^{N_\nu} + \int_{N_\nu}^{n_{\nu+1}/2} + \int_{n_{\nu+1}/2}^{r_1-s_1}\right]\left|\frac{d}{dx}\left\{\frac{c(x)}{e^{i\gamma(x)} - 1}\right\}\right|dx$$
$$= 2(I_1 + I_2 + I_3 + I_4).$$

Using (1.5) and the fact that $\gamma'(x) = O(1/x)$ we have

* In case $2n_\nu > N_\nu$ we drop I_2, if $N_\nu > \frac{1}{2}n_{\nu+1}$, we drop I_3, and if $\frac{1}{2}n_{\nu+1} > n_{\nu+1} - n_{\nu+1}^{3/5}$, we drop I_3. It is essential that $n_\nu + n_\nu^{3/5} \leqq n_{\nu+1} - n_{\nu+1}^{3/5}$. This will be shown later (cf. (2.1)).

$$I_1 = O\left\{\int_{r+s}^{2n_\nu} \frac{x^{1+\epsilon}\,|\,c'(x)\,|}{x-n_\nu}\,dx + \int_{r+s}^{2n_\nu} \frac{c(x)x^{1+2\epsilon}}{(x-n_\nu)^2}\,dx\right\}$$
$$= O\{I_1' + I_1''\},$$

and by (B) and (C)

$$(1.9)\qquad \begin{aligned} I_1' &= O\left\{\int_s^{n_\nu} [(1-\rho)b(n_\nu+x)(n_\nu+x)^{1+\epsilon}\rho^{x+n_\nu} \right.\\ &\qquad\qquad \left. + b(x+n_\nu)(x+n_\nu)^{\delta+\epsilon}\rho^{x+n_\nu}]\,\frac{dx}{x}\right\} \\ &= O\{[(1-\rho)b(n_\nu)n_\nu^{1+\epsilon}\rho^{n_\nu} + b(n_\nu)n_\nu^{\delta+\epsilon}\rho^{n_\nu}]\log n_\nu\} \end{aligned}$$

and

$$(1.10)\qquad I_1'' = O\left\{\int_s^{n_\nu} c(n_\nu+x)(n_\nu+x)^{1+2\epsilon}x^{-2}dx\right\} = O\{c(n_\nu)n_\nu^{2/5+2\epsilon}\}.$$

Similarly

$$I_2 = O(I_2' + I_2''),$$

where

$$(1.11)\qquad \begin{aligned} I_2' &= \int_{2n_\nu}^{N_\nu} x^\epsilon[(1-\rho)b(x)\rho^x + x^{\delta-1}b(x)\rho^x]dx \\ &= O\left\{\sum_{2n_\nu}^{N_\nu} n^\epsilon[(1-\rho)b(n)\rho^n + n^{\delta-1}b(n)\rho^n]\right\} \end{aligned}$$

and

$$(1.12)\qquad I_2'' = O\left\{\sum_{2n_\nu}^{N_\nu} c(n)n^{2\epsilon-1}\right\}.$$

Also $I_3 = O(I_3' + I_3'')$, and by (1.7), if $x \geqq N_\nu$, $2^{-j-2}n_{\nu+1} \leqq x \leqq 2^{-j-1}n_{\nu+1}$,

$$(e^{i\gamma(x)} - 1)^{-1} = O\{(n_{\nu+1}2^{-j})^{1+\epsilon}(n_{\nu+1}2^{-j} - x)^{-1}\} = O\{x^\epsilon\}$$

uniformly for $j = 0, 1, \cdots$, so that

$$(1.13)\qquad I_3' = O\left\{\sum_{N_\nu}^{n_{\nu+1}/2} n^\epsilon[(1-\rho)b(n)\rho^n + n^{\delta-1}b(n)\rho^n]\right\},$$

$$(1.14)\qquad I_3'' = O\left\{\sum_{N_\nu}^{n_{\nu+1}/2} n^{2\epsilon-1}b(n)\rho^n\right\}.$$

Finally $I_4 = O(I_4' + I_4'')$ and by (1.6)

$$
\begin{aligned}
I_4' &= \int_{n_{\nu+1/2}}^{r_1-s_1} [(1-\rho)b(x)n_{\nu+1}^{1+\epsilon}\rho^x + b(x)x^{\delta-1}n_{\nu+1}^{1+\epsilon}\rho^x](n_{\nu+1}-x)^{-1}dx \\
(1.15)\qquad &= O\{n_{\nu+1}^{1+\epsilon}\log n_{\nu+1}[(1-\rho)b(n_{\nu+1})\rho^{n_{\nu+1/2}} + b(n_{\nu+1})n_{\nu+1}^{\delta-1}\rho^{n_{\nu+1/2}}]\},
\end{aligned}
$$

$$
(1.16)\quad I_4'' = O\{n_{\nu+1}^{2/5+2\epsilon}\, b(n_{\nu+1})\rho^{n_{\nu+1/2}}\}.
$$

From (1.4) and (1.8)–(1.16) we get the principal result of the paper that

$$
f(z) = \sum_{n=1}^{\infty} b(n)e^{i\Delta(n)}z^n = G(z) + R(z),
$$

where

$$
G(z) = \pi^{1/2}e^{\pi i/4}\sum_{\nu=1}^{\infty} b(n_\nu)[d'(n_\nu)]^{-1/2}e^{iF(\nu;\theta)}\rho^{n_\nu}
$$

and, if $\alpha = \max(\eta, 2\epsilon)$,

$$
\begin{aligned}
R(z) = O\Big\{ \sum_{\nu=1}^{\infty} &[(1-\rho)b(n_\nu)n_\nu^{6/5+\epsilon}\rho^{n_\nu} + b(n_\nu)n_\nu^{2/5+\alpha}\rho^{n_\nu} \\
&+ (1-\rho)b(n_\nu)n_\nu^{6/5+\epsilon}\rho^{n_\nu/2} + b(n_\nu)n_\nu^{2/5+\alpha}\rho^{n_\nu/2}] \\
&+ \sum_{n=1}^{\infty}[b(n)n^{2\epsilon-1}\rho^n + (1-\rho)b(n)n^\epsilon\rho^n]\Big\} = O\Big\{\sum_{i=1}^{6} R_i(z)\Big\}.
\end{aligned}
$$

We notice that this procedure is valid only for $|z| < 1$. We shall now give a functional equation of the type used by Wilton (5) which is valid for $|z| = 1$. The equation is

$$
(1.17)\quad \sum_{n=1}^{m} b(n)e^{i\Delta(n)}e^{in\theta} = \pi^{1/2}e^{\pi i/4}\sum_{\nu=1}^{\mu} b(n_\nu)[d'(n_\nu)]^{-1/2}e^{iF(\nu;\theta)} + R(m, \theta),
$$

where $n_\mu + n_\mu^{3/5} \leqq m \leqq n_{\mu+1} + n_{\mu+1}^{3/5}$. It can be seen that the sums

$$
\sum_{r+s}^{r_1+s_1} b(n)e^{i\Delta(n)}e^{in\theta}, \qquad \nu < \mu,
$$

$$
\sum_{u+v}^{m} b(n)e^{i\mu(n)}e^{in\theta}, \qquad m \leqq n_{\mu+1} - n_{\mu+1}^{3/5},
$$

where $u = n_\mu$, $v = n_\mu^{3/5}$, can be handled by the methods which we have used and moreover

$$
\sum_{u_1-v_1}^{m} b(n)e^{i\Delta(n)}e^{in\theta} = O\{b(m)[d'(m)]^{-1/2}\},
$$

where $u_1 = n_{\mu+1}$, $v_1 = n_{\mu+1}^{3/5}$. This shows that

$$R(m, \theta) = O\left\{ \sum_{\nu=1}^{\mu} b(n_\nu)n_\nu^{2/5+\alpha} + \sum_{n=1}^{m} b(n)n^{2\epsilon-1} + b(m)[d'(m)]^{-1/2} \right\}.$$

2. **The order of** $G(z)$. If the series

$$\sum_{\nu=1}^{\infty} b(n_\nu)[d'(n_\nu)]^{-1/2}$$

converges uniformly in θ, then the function $G(z)$ tends to a continuous function of θ as $\rho = |z| \to 1$. We shall now suppose that the series does not converge and investigate the order of $G(z)$ as $\rho \to 1$. We notice some properties of $n_\nu(\theta)$. First, since $d(x)$ is monotone, $n_\nu(\theta) \uparrow \infty$ as $\nu \to \infty$ and $n_\nu(\theta)$ is monotone decreasing as a function of θ on $(0, 2\pi)$. Also

$$n_\nu(0) = n_{\nu+1}(2\pi).$$

Finally, since

$$d\{n_\nu(\theta)\} = 2\pi\nu - \theta,$$

we have

$$d'\{n_\nu(\theta)\} \frac{d}{d\nu} n_\nu(\theta) = 2\pi$$

and, since $d'(x) = O(1/x)$, there must be a constant $c > 0$ such that

$$(2.1) \qquad n_{\nu+1}(\theta) = n_\nu(\theta) + \int_\nu^{\nu+1} \frac{d}{d\mu} n_\mu(\theta) d\mu > n_\nu(\theta) + cn_\nu(\theta) = (1 + c)n_\nu(\theta).$$

We let $h(x) = b(x)\ [d'(x)]^{-1/2}$ and, if $n_\mu(0) \leqq x < n_{\mu+1}(0)$, we define

$$M(x) = \sum_{\nu=2}^{\mu} \max_{n_{\nu-1}(0) \leqq y \leqq n_\nu(0)} h(y) + \max_{n_\mu(0) \leqq y \leqq x} h(y).$$

If A is greater than 1 and $n_\mu(0) \leqq x < n_{\mu+1}(0)$, $n_{\mu'}(0) \leqq Ax < n_{\mu'+1}(0)$, then by (2.1) we have

$$\mu' - \mu < \log A\,[\log(1 + c)]^{-1},$$

and by (C), $h(y) = O\{h(x)\}$, $x \leqq y \leqq Ax$, so that, if $n_\mu(0) \leqq x \leqq n_{\mu+1}(0)$, then

$$(2.2) \qquad M(Ax) = \sum_{\nu=2}^{\mu} \max_{n_{\nu-1}(0) \leqq y \leqq n_\nu(0)} h(y) + O\left\{ \max_{n_\mu(0) \leqq y \leqq x} h(y) \right\} = O\{M(x)\}.$$

We write

$$|G(z)| \leqq 2\pi^{1/2} \sum_{\nu=1}^{\infty} h(n_\nu)\rho^{n_\nu} = 2\pi^{1/2} \sum_{\nu=1}^{\infty} h(n_\nu)e^{yn_\nu}, \qquad y = \log \rho,$$

and define μ so that $n_\mu(\theta) \leqq -1/y < n_{\mu+1}(\theta)$. Then

$$\sum_{\nu=1}^{\infty} h(n_\nu)e^{yn_\nu} = \sum_{\nu=1}^{\mu} h(n_\nu)e^{yn_\nu} + \sum_{\nu=\mu+1}^{\infty} h(n_\nu)e^{yn_\nu}$$

and, if $0 \leqq \theta < 2\pi$,

$$\sum_{\nu=1}^{\mu} h(n_\nu)e^{yn_\nu} = O\left\{ \sum_{\nu=1}^{\mu} h(n_\nu) \right\} = O\{ M[n_\mu(\theta)] \}.$$

By (C)

$$\sum_{\nu=\mu+1}^{\infty} h(n_\nu)e^{yn_\nu} = h\left(-\frac{1}{y}\right) \sum_{\nu=\mu+1}^{\infty} h(n_\nu) \left[h\left(-\frac{1}{y}\right)\right]^{-1} e^{yn_\nu}$$

$$= O\left\{ h\left(-\frac{1}{y}\right) \sum_{\nu=\mu+1}^{\infty} (-n_\nu y)^{\log k} e^{n_\nu y} \right\}.$$

Since $-n_\nu y > (1+c)^{\nu-\mu}(-n_\mu y) > (1+c)^{\nu-\mu-1}$, we may find an A and an m not depending on y so that

$$n_\nu y + \log k \log(-n_\nu y) < -A(\nu - \mu), \text{ for } (\nu - \mu) > m.$$

This shows that

$$\sum_{\nu=\mu+1}^{\infty} h(n_\nu)e^{yn_\nu} = O\left\{ h\left(-\frac{1}{y}\right) \sum_{\mu=\nu+1}^{\infty} \exp\left[n_\nu y + \log k \log(-n_\nu y)\right] \right\}$$

$$= O\left\{ h\left(-\frac{1}{y}\right)\right\}.$$

Hence by (2.2)

$$G(z) = O\left\{ M\left(\frac{1}{1-\rho}\right)\right\}. \tag{2.3}$$

We shall now use the additional hypothesis

$$M(x) = O\{ h(x) \} \tag{D}$$

and we propose to prove that if (D) is satisfied then

$$G(z) = \Omega\left\{ h\left(\frac{1}{1-\rho}\right)\right\} = \Omega\left\{ M\left(\frac{1}{1-\rho}\right)\right\}.$$

The method is essentially that used by Hardy [1] for a similar problem. We take the series

$$\sum_{\nu=1}^{\infty} n_\nu^\alpha h(n_\nu) e^{yn_\nu} e^{iF(\nu;\theta)}$$

and let $y=-\alpha/n_\mu$. Then

$$\left|\sum_{\nu=1}^{\infty} n_\nu^\alpha h(n_\nu)e^{yn_\nu}e^{iF(\nu;\theta)}\right| \geqq n_\mu^\alpha h(n_\nu)e^{-\alpha} - \sum_{\nu=1}^{\mu-1} n_\nu^\alpha h(n_\nu)e^{yn_\nu} - \sum_{\nu=\mu+1}^{\infty} n_\nu^\alpha h(n_\nu)e^{yn_\nu}.$$

By (D) we have $h(n_\nu)=O\{h(n_\mu)\}$, $\nu<\mu$, and hence there exists a $\lambda>0$ and independent of α, ν, and λ, such that

$$\left(\frac{n_\nu}{n_\mu}\right)^\alpha \frac{h(n_\nu)}{h(n_\mu)} < \lambda\left(\frac{n_\nu}{n_\mu}\right)^\alpha.$$

Therefore

$$\begin{aligned}\sum_{\nu=1}^{\mu-1} n_\nu^\alpha h(n_\nu)\exp\left[-\alpha n_\nu/n_\mu\right] &< \lambda h(n_\mu)n_\mu^\alpha e^{-\alpha}\sum_{\nu=1}^{\mu-1}\left(\frac{n_\nu}{n_\mu}\right)^\alpha \exp\left[-\alpha\left(\frac{n_\nu}{n_\mu}-1\right)\right] \\ &= \lambda h(n_\mu)n_\mu^\alpha e^{-\alpha}\sum_{\nu=1}^{\mu-1}\exp\left[-\alpha\left(\frac{n_\nu}{n_\mu}-1-\log\frac{n_\nu}{n_\mu}\right)\right].\end{aligned}$$

But, since $n_\nu/n_\mu<(1+c)^{-1}$, for $\nu<\mu$, there exists a $p>0$ such that

$$\frac{n_\nu}{n_\mu}-1-\log\frac{n_\nu}{n_\mu}>p, \quad \text{for} \quad \nu<\mu,$$

and, for $\mu-\nu>u$, u depending only on c, there exists an A so that

$$\frac{n_\nu}{n_\mu}-1-\log\frac{n_\nu}{n_\mu}>A\left[\log(1+c)\right](\mu-\nu).$$

Therefore

$$\begin{aligned}\sum_{\nu=1}^{\mu-1}\left(\frac{n_\nu}{n_\mu}\right)^\alpha \exp\left[-\alpha\left(\frac{n_\nu}{n_\mu}-1\right)\right] &< \sum_{\nu=1}^{\mu-u-1}\exp\left[-\alpha A(\mu-\nu)\log(1+c)\right] \\ + \sum_{\nu=\mu-u}^{\mu-1} e^{-\alpha p} &< \sum_{\nu=1}^{\infty}\exp\left[-\alpha A\nu\log(1+c)\right]+ue^{-\alpha p}=o(1) \text{ as } \alpha\to\infty.\end{aligned}$$

We have also by (C)

$$\begin{aligned}&\sum_{\nu=\mu+1}^{\infty} n_\nu^\alpha h(n_\nu)\exp\left(-\alpha n_\nu/n_\mu\right) \\ &\quad = O\left\{n_\mu^\alpha h(n_\mu)e^{-\alpha}\sum_{\nu=\mu+1}^{\infty}\left(\frac{n_\nu}{n_\mu}\right)^{\alpha+\log k}\exp\left[-\alpha\left(\frac{n_\nu}{n_\mu}-1\right)\right]\right\} \\ &\quad = O\left\{n_\mu^\alpha h(n_\mu)e^{-\alpha}\sum_{\nu=\mu+1}^{\infty}\left[-\alpha\left(\frac{n_\nu}{n_\mu}-1-\log\frac{n_\nu}{n_\mu}\right)+\log k\log\frac{n_\nu}{n_\mu}\right]\right\}\end{aligned}$$

and, since $n_\nu/n_\mu>(1+c)$ for $\nu>\mu$, there exists a $p>0$ for which,

$$\frac{n_\nu}{n_\mu} - 1 - \log \frac{n_\nu}{n_\mu} > p,$$

and hence for α greater than a suitable α_0, there is a $p_0 > 0$, such that

$$\alpha\left(\frac{n_\nu}{n_\mu} - 1 - \log \frac{n_\nu}{n_\mu}\right) - \log k \log \frac{n_\nu}{n_\mu} > \alpha p_0.$$

For $\nu - \mu > w$, w independent of α for $\alpha > \alpha_0$,

$$\alpha\left(\frac{n_\nu}{n_\mu} - 1 - \log \frac{n_\nu}{n_\mu}\right) + \log k \log \frac{n_\nu}{n_\mu} > \alpha[\log (1 + c)](\nu - \mu).$$

Hence

$$\sum_{\nu=\mu+1}^{\infty} \exp \left[- \alpha\left(\frac{n_\nu}{n_\mu} - 1 - \log \frac{n_\nu}{n_\mu}\right) - \log k \log \frac{n_\nu}{n_\mu}\right]$$
$$< we^{-\alpha p_0} + \sum_{\nu=1}^{\infty} \exp [- \alpha\nu \log (1 + c)] = o(1) \text{ as } \alpha \to \infty.$$

Consequently we can find an α_1 so that for $\alpha > \alpha_1$,

$$\sum_{\nu=1}^{\mu-1} n_\nu^\alpha h(n_\nu) \exp (- \alpha n_\nu / n_\mu) + \sum_{\nu=\mu+1}^{\infty} n_\nu^\alpha h(n_\nu) \exp (- \alpha n_\nu / n_\mu) < \tfrac{1}{2} n_\mu^\alpha h(n_\mu) e^{-\alpha}$$

uniformly in θ and μ, and by (C), for $y = -(\alpha/n_\mu)$,

$$\left| \frac{d^\alpha}{d\rho^\alpha} G(z) \right| > \tfrac{1}{2}\pi^{1/2} n_\mu^\alpha h(n_\mu) e^{-\alpha} > \omega\left(-\frac{1}{y}\right)^\alpha h\left(-\frac{1}{y}\right), \quad \omega \text{ depends on } \alpha.$$

We can easily see that

$$\frac{d^\alpha}{d\rho^\alpha} G(z) = O\left\{\left(-\frac{1}{y}\right)^\alpha h\left(-\frac{1}{y}\right)\right\}.$$

Therefore by a theorem of Hardy and Littlewood* we must have

$$G(z) = \Omega\left\{h\left(\frac{1}{1 - \rho}\right)\right\}.$$

Therefore under condition (D) we have an exact characterization of the order of $G(z)$ as $|z| \to 1$.

3. **The order of** $R(z)$. By the above method we can easily show that

$$\sum_{\nu=1}^{\infty} b(n_\nu) n_\nu^{2/5+\alpha} \rho^{n_\nu} = O\left\{K_1\left(\frac{1}{1 - \rho}\right)\right\}, \tag{3.1}$$

* Hardy and Littlewood [3]. This result follows from Theorem 8 on setting $\phi = \psi = h(-1/y)$.

where, if $n_\mu(0) \leqq x < n_{\mu+1}(0)$,

$$K_1(x) = \sum_{\nu=2}^{\mu} \max_{n_{\nu-1}(0) \leqq y \leqq n_\nu(0)} b(y) y^{2/5+\alpha} + \max_{n_\mu(0) \leqq y < x} b(y) y^{2/5+\alpha}.$$

Also

$$\sum_{\nu=1}^{\infty} b(n_\nu) n_\nu^{6/5+\epsilon} \rho^{n_\nu} = O\left\{K_2\left(\frac{1}{1-\rho}\right)\right\}, \tag{3.2}$$

where, if $n_\mu(0) \leqq x < n_{\mu+1}(0)$,

$$K_2(x) = \sum_{\nu=2}^{\mu} \max_{n_{\nu-1}(0) \leqq y \leqq n_\nu(0)} b(y) y^{6/5+\epsilon} + \max_{n_\mu(0) \leqq y \leqq x} b(y) y^{6/5+\epsilon} = O\{x^{4/5} K_1(x)\}.$$

Similarly we can show that

$$\sum_{\nu=1}^{\infty} b(n_\nu) n_\nu^{2/5+\alpha} \rho^{n_\nu/2} = O\left\{K_1\left(\frac{2}{1-\rho}\right)\right\} = O\left\{K_1\left(\frac{1}{1-\rho}\right)\right\} \tag{3.3}$$

and

$$\sum_{\nu=1}^{\infty} b(n_\nu) n_\nu^{6/5+\epsilon} \rho^{n_\nu/2} = O\left\{K_2\left(\frac{2}{1-\rho}\right)\right\} = O\left\{(1-\rho)^{-4/5} K_1\left(\frac{1}{1-\rho}\right)\right\}. \tag{3.4}$$

Since $2/5+\alpha < \frac{1}{2}$, we have

$$x^{2/5+\alpha} = o\{[d'(x)]^{-1/2}\}$$

and hence, if $M(x) \to \infty$ as $x \to \infty$,

$$K_1(x) = o\{M(x)\}.$$

Therefore, if $M(x) \to \infty$ as $x \to \infty$,

$$R_j(z) = o\left\{M\left(\frac{1}{1-\rho}\right)\right\}, \qquad j = 1, 2, 3, 4. \tag{3.5}$$

If $N \leqq -1/y < N+1$, $y = \log \rho$, by (C)

$$\sum_{n=N+1}^{\infty} b(n) n^{2\epsilon-1} e^{yn} = O\left\{\sum_{j=0}^{\infty} \frac{b(2^j N)}{(2^j N)^{1-2\epsilon}} \sum_{n=2^j N+1}^{2^{j+1}N} e^{-n/N}\right\}$$

and, since

$$b(2^j N)(2^j N)^{2\epsilon-1} = O\{2^{j \log k} b(N)(2^j N)^{2\epsilon-1}\},$$

we have

$$\sum_{n=N+1}^{\infty} b(n)n^{2\epsilon-1}\rho^n = O\left\{b(N)N^{2\epsilon-1}\sum_{j=0}^{\infty} 2^{j(\log k+2\epsilon-1)}e^{-2^j}\right\}$$
$$= O\{b(N)N^{2\epsilon-1}\}.$$

Therefore

$$R_5(z) = O\left\{\sum_{n=1}^{[1/(1-\rho)]} b(n)n^{2\epsilon-1}\right\}. \tag{3.6}$$

Similarly

$$R_6(z) = O\left\{(1-\rho)\sum_{n=1}^{[1/(1-\rho)]} b(n)n^{\epsilon}\right\} = O\left\{\sum_{n=1}^{[1/(1-\rho)]} b(n)n^{\epsilon-1}\right\}. \tag{3.7}$$

If the series

$$\sum_{\nu=1}^{\infty} b(n_\nu)[d'(n_\nu)]^{-1/2}$$

converges uniformly in θ, it is easily seen from the above estimates that the functions $R_i(z)$ tend to a continuous function of θ as $\rho\to 1$, for $i=1, 2, 3, 4$. Moreover, since $d'(x)=O(1/x)$ we must have $b(x)=O(x^{-1/2})$ so that

$$\sum_{n=1}^{\infty} b(n)n^{2\epsilon-1}$$

converges and

$$\sum_{n=1}^{\infty} b(n)n^{\epsilon}\rho^n = o\{(1-\rho)^{-1}\}.$$

Therefore in this case $R(z)$ is continuous for $|z|=1$.

We now wish to give some estimate for

$$\sum_{n=1}^{x} b(n)n^{2\epsilon-1}$$

and compare it if possible with $M(x)$. If $b(x)$ is increasing, by the monotonicity of $b(x)$

$$\sum_{n=1}^{N} b(n)n^{2\epsilon-1} = O\{b(N)N^{2\epsilon}\}$$

and, since by (A)

$$b(N)N^{1/2} = O\{M(N)\},$$

we see that

$$R_j(z) = o\left\{M\left(\frac{1}{1-\rho}\right)\right\}, \qquad j = 5, 6.$$

If $b(x)$ is not increasing the problem might have to be considered differently in separate cases. However if we make the reasonable assumption that

$$\sum_{n=1}^{x} b(n)n^{2\epsilon-1} = O\{b(x)x^{2\epsilon+\gamma}\}, \qquad 2\epsilon + \gamma < \tfrac{1}{2},$$

then we see that in this case

$$R_j(z) = o\left\{M\left(\frac{1}{1-\rho}\right)\right\}.$$

This assumption is always satisfied for functions of the logarithmic-exponential type which is the type of function used in applications of this method.

4. **The modulus of continuity.** If a function $f(z)$ analytic for $|z|<1$ is continuous for $|z|=1$, we define the modulus of continuity of $f(z)$ as the function

$$\omega(f, h) = \underset{|\theta_1-\theta_2|\leqq h}{\text{l.u.b.}} |f(e^{i\theta_1}) - f(e^{i\theta_2})|.$$

We wish to develop a method of finding the order of $\omega(f, h)$. Let us suppose that there is a function $g(x)$ such that $g(x)\uparrow\infty$ as $x\to 1$, then we shall show that if

$$f'(z) = O\{g(\rho)\},$$

then

$$\omega(f, h) = O\left\{\int_{1-h}^{1} g(x)dx\right\},$$

provided the latter integral exists. The method used has been used by Hardy and Littlewood [4].

We have

$$|f(e^{i\theta_1}) - f(e^{i\theta_2})| = \lim_{\rho\to 1} |f(\rho e^{i\theta_1}) - f(\rho e^{i\theta_2})|,$$

so that we need only show that

$$|f(\rho e^{i\theta_1}) - f(\rho e^{i\theta_2})| = O\left\{\int_{1-h}^{1} g(x)dx\right\}$$

uniformly in $\rho<1$ and $|\theta_1-\theta_2|\leqq h$. Since $f'(z)$ is analytic for $|z|<1$,

$$f(\rho e^{i\theta_1}) - f(\rho e^{i\theta_2}) = \int_{\rho e^{i\theta_2}}^{\rho e^{i\theta_1}} f'(z)dz = \left\{\int_1 + \int_2 + \int_3\right\} f'(z)dz,$$

where $\int_1$ is the integral taken along the radius $\theta=\theta_2$ from ρ to $\rho-h$, and $\int_3$

is the similar integral along $\theta=\theta_1$ from $\rho-h$ to ρ, while $\int_2$ is the integral along the arc $|z|=\rho-h$ from θ_2 to θ_1. Then

$$\int_2 f'(z)dz = i(\rho - h)\int_{\theta_2}^{\theta_1} e^{i\theta}f'[(\rho - h)e^{i\theta}]d\theta$$
$$= O\{\,|\,\theta_1 - \theta_2|g(\rho - h)\} = O\{hg(1 - h)\}.$$

Since $g(x)\uparrow\infty$ as $x\to 1$,

$$\int_1 f'(z)dz = \int_\rho^{\rho-h} e^{i\theta_2}f'(xe^{i\theta_2})dx$$
$$= O\left\{\int_{\rho-h}^{\rho} g(x)dx\right\}$$
$$= O\left\{\int_{1-h}^{1} g(x)dx\right\}.$$

Similarly

$$\int_3 f'(z)dz = O\left\{\int_{1-h}^{1} g(x)dx\right\}.$$

Finally we notice that, since $g(x)\uparrow\infty$ as $x\to 1$,

$$hg(1 - h) \leqq \int_{1-h}^{1} g(x)dx.$$

Therefore

$$\omega(f, h) = O\left\{\int_{1-h}^{1} g(x)dx\right\}.$$

5. **Applications.** To illustrate the scope of this method we shall apply it in detail to a particular function which has already been considered and point out some extensions of the usual results which may be obtained by this method. The function which we shall consider is

$$f_{\beta\gamma}(z) = \sum_{n=2}^{\infty} n^{-1/2}(\log n)^{-\gamma}\exp\left[i\int_1^n(\log x)^\beta dx\right]z^n, \qquad 0 \leqq \beta \leqq 1.$$

This is similar to the function

$$\sum_{n=2}^{\infty} n^{-1/2}(\log n)^{-\gamma}\exp\,[in(\log n)^\beta]z^n$$

considered by Ingham [6]. Ingham's function could also be handled by these methods but the details would be slightly more complicated.

For the function under consideration

$$d(x) = (\log x)^\beta, \qquad b(x) = x^{-1/2}(\log x)^{-\gamma}$$

and

$$(\log [n_\nu(\theta)])^\beta = 2\pi\nu - \theta, \qquad n_\nu(\theta) = \exp\left[(2\pi\nu - \theta)^{1/\beta}\right].$$

Therefore

$$d'(x) = \beta x^{-1}(\log x)^{\beta-1} = O(x^{-1}), \quad \text{since} \quad \beta \leqq 1,$$

$$\text{(A)} \qquad \frac{1}{d'(x)} = \frac{1}{\beta}\, x(\log x)^{\beta-1} = O(x^{1+\epsilon}), \quad \text{for every } \epsilon,$$

$$d''(x) = O(x^{-2});$$

$$\text{(B)} \qquad \begin{aligned} b'(x) &= -\tfrac{1}{2}x^{-3/2}(\log x)^{-\gamma} - \gamma x^{-3/2}(\log x)^{-\gamma-1} = O\{x^{-3/2}(\log x)^{-\gamma}\} \\ &= O\{b(x)x^{-1}\}; \end{aligned}$$

$$\text{(C)} \qquad \begin{aligned} e^2 d'(x') &> d'(x), \qquad x \leqq x' \leqq ex, \\ e^2 b(x') &> b(x), \qquad x \leqq x' \leqq ex. \end{aligned}$$

Since $\eta = \delta = 0$ and ϵ can be made less than $1/30$, we have

$$\begin{aligned} R(z) = O\Bigg\{&(1-\rho)\sum_{\nu=2}^{\infty} n_\nu^{7/10}\rho^{n_\nu} + \sum_{\nu=2}^{\infty} n_\nu^{-1/20}\rho^{n_\nu} \\ &+ (1-\rho)\sum_{n=1}^{\infty} n^{-3/5}\rho^{n} + \sum_{n=1}^{\infty} n^{-6/5}\rho^{n}\Bigg\} \\ = O\{&(1-\rho)(1-\rho)^{-7/10} + 1 + (1-\rho)(1-\rho)^{-2/5} + 1\}, \end{aligned}$$

so that $R(z)$ is continuous when $|z| = 1$. The function $G(z)$ is continuous for $|z| = 1$ if

$$\text{(5.1)} \qquad \sum_{\nu=2}^{\infty} n_\nu^{-1/2}(\log n_\nu)^{-\gamma} n_\nu^{1/2}(\log n_\nu)^{(1-\beta)/2} = \sum_{\nu=2}^{\infty} (2\pi\nu - \theta)^{-\gamma/\beta - 1/2 + 1/(2\beta)}$$

converges, and the necessary and sufficient condition that (5.1) converge is

$$\frac{\gamma}{\beta} + \frac{1}{2} - \frac{1}{2\beta} > 1, \qquad \text{or} \qquad \gamma > \frac{1}{2}(\beta + 1).$$

This corresponds to the result obtained by Ingham.

We shall now consider the modulus of continuity of $f_{\beta\gamma}(z)$. We notice that the function $b(x) = x^{1/2}(\log x)^{-\gamma}$ satisfies (B) and (C) so that the discussion of §§2, 3 applies to the function

$$f'_{\beta\gamma}(z) = \sum_{n=1}^{\infty} n^{1/2}(\log n)^{-\gamma} \exp\left\{i\int_1^n (\log x)^\beta dx\right\} z^n.$$

Since $\beta \leqq 1$,

$$M(x) = \sum_{n_\nu \leqq x} n_{\nu-1}(\log n_{\nu-1})^{-\gamma-\beta/2+1/2} + x(\log x)^{-\gamma-\beta/2+1/2}$$
$$= O\{x(\log x)^{-\gamma-\beta/2+1/2}\},$$

so that

$$f'_{\beta\gamma}(z) = O\left\{(1-\rho)^{-1}\left[\log\left(\frac{1}{1-\rho}\right)\right]^{-\gamma-\beta/2+1/2}\right\}.$$

But

$$\int_{1-h}^{1}(1-\rho)^{-1}\left[\log\left(\frac{1}{1-\rho}\right)\right]^{-\gamma-\beta/2+1/2} d\rho = \int_0^h x^{-1}\left(\log\frac{1}{x}\right)^{-\gamma-\beta/2+1/2} dx$$
$$= \int_{1/h}^{\infty} x^{-1}(\log x)^{-\gamma-\beta/2+1/2}dx = \frac{-1}{\gamma+\beta/2-3/2}\left(\log\frac{1}{h}\right)^{-\gamma-\beta/2+3/2},$$
$$\text{if } \gamma > \frac{3}{2}-\frac{\beta}{2}$$

and hence

$$\omega(f_{\beta\gamma}, h) = O\left\{\left(\log\frac{1}{h}\right)^{-\gamma-\beta/2+3/2}\right\}, \qquad \text{if} \qquad \gamma > \frac{3}{2}-\frac{\beta}{2}\cdot$$

This result could not be obtained by the other methods to which we have referred.

6. **Further applications.** If

$$\sum_{n=1}^{\infty} b^2(n) = \infty,$$

we know that the function

$$f(z) = \sum_{n=1}^{\infty} b(n)e^{i\Delta(n)}z^n$$

can not tend to a continuous or even a bounded function as $\rho = |z| \to 1$. We wish now to study the behavior of $f(\rho e^{i\theta})$ for a particular choice of $d(x)$. We assume that $b(x)$ satisfies (B) with $\delta < 1/10 - \epsilon$, and (C) and,

$$\text{(E)} \qquad \int_1^x b^2(t)dt = O\{x^{1+\epsilon}b^2(x)\}, \qquad \epsilon < \frac{1}{30}\cdot$$

We define

$$d(x) = \log\left[\int_1^x b^2(t)dt\right].$$

Then by hypothesis $d(x) \uparrow \infty$, as $x \to \infty$, and

$$d'(x) = \frac{b^2(x)}{\int_1^x b^2(t)dt} = O(x^{-1}),$$

$$d''(x) = O(x^{\delta-2}), \qquad \delta < \frac{1}{10} - \epsilon,$$

$$\frac{1}{d'(x)} = O(x^{1+\epsilon}), \qquad \epsilon < \frac{1}{30},$$

so that the function $d(x)$ so defined satisfies (A). It is also clear that since $b(x)$ satisfies (C), $d(x)$ will also satisfy (C). We have

$$b(x)[d'(x)]^{-1/2} = \left(\int_1^x b^2(t)dt\right)^{1/2} \uparrow \infty \quad \text{as} \quad x \to \infty$$

and

$$b(n_\nu)[d'(n_\nu)]^{-1/2} = e^{\pi\nu - \theta/2}.$$

Therefore

$$\begin{aligned} M(x) &= \sum_{n_\nu(0) \leqq x} b(n_\nu(0))[d'(n_\nu(0))]^{-1/2} + b(x)[d'(x)]^{-1/2} \\ &= \sum_{n_\nu(0) \leqq x} e^{\pi\nu} + b(x)[d'(x)]^{-1/2} = O\{b(x)[d'(x)]^{-1/2}\} \end{aligned}$$

and condition (D) is satisfied. By the argument of §3

$$R(z) = o\left\{M\left(\frac{1}{1-\rho}\right)\right\} + O\left\{\sum_1^{[1/(1-\rho)]} b(n)n^{2\epsilon-1}\right\}$$

and

$$\begin{aligned} \sum_1^N b(n)n^{2\epsilon-1} &\leqq \left[\sum_1^N n^{-1/2}b^2(n)\right]^{1/2}\left[\sum_1^N n^{4\epsilon-3/2}\right]^{1/2} \\ &= o\{M(N)\}, \end{aligned}$$

so that

$$R(z) = o\left\{M\left(\frac{1}{1-\rho}\right)\right\}.$$

Therefore by the argument of §2,

$$f(z) = O\left\{M\left(\frac{1}{1-\rho}\right)\right\} = O\left\{\left(\int_1^{1/(1-\rho)} b^2(x)dx\right)^{1/2}\right\}$$

and

$$f(z) = \Omega\left\{M\left(\frac{1}{1-\rho}\right)\right\}.$$

This answers the question mentioned in the introduction about the possibility of choosing $d(x)$ so that $f(z)$ have the same order as $|z|\to 1$ for every θ.

We might apply these considerations to the case where $b(x)=x^{-1/2}$. We can readily obtain the result for this case that if

$$d(x) = \log\left[\int_1^x \frac{dt}{t}\right] = \log\log x$$

then

$$f(z) = \sum_{n=3}^{\infty} n^{-1/2}\exp\left[i\Delta(n)\right]z^n = O\left\{\left[\log\left(\frac{1}{1-\rho}\right)\right]^{1/2}\right\}$$

and

$$f(z) = \Omega\left\{\left[\log\left(\frac{1}{1-\rho}\right)\right]^{1/2}\right\}.$$

We might compare this with the results for the function

$$f_0(z) = \sum_{n=2}^{\infty} n^{-1/2}\exp\left[in\log n\right]z^n$$

obtained by Hardy and Littlewood which are mentioned in the introduction ((1) and (2)).

References

1. Hardy, these Transactions, vol. 17 (1916), pp. 301–325.
2. Hardy and Littlewood, Proceedings of the National Academy of Sciences, vol. 2 (1916), pp. 583–586.
3. Hardy and Littlewood, Proceedings of the London Mathematical Society, (2), vol. 11 (1912), pp. 411–478.
4. Hardy and Littlewood, Mathematische Zeitschrift, vol. 28 (1928), pp. 612–634.
5. Hille, Journal of the London Mathematical Society, vol. 4 (1929), pp. 176–182.
6. Ingham, Annals of Mathematics, (2), vol. 31 (1930), pp. 241–250.
7. Paley, Proceedings of the London Mathematical Society, (2), vol. 31 (1930), pp. 301–328.
8. Wilton, Journal of the London Mathematical Society, vol. 9 (1934), pp. 194–201, 247–256.

Northwestern University,
Evanston, Ill.

ON THE SOLUTIONS OF QUASI-LINEAR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS*

BY

CHARLES B. MORREY, JR.

In this paper, we are concerned with the existence and differentiability properties of the solutions of "quasi-linear" elliptic partial differential equations in two variables, i.e., equations of the form

$$A(x, y, z, p, q)r + 2B(x, y, z, p, q)s + C(x, y, z, p, q)t = D(x, y, z, p, q)$$

$$AC - B^2 > 0, A > 0 \left(p = \frac{\partial z}{\partial x}, \; q = \frac{\partial z}{\partial y}, \; r = \frac{\partial^2 z}{\partial x^2}, \; s = \frac{\partial^2 z}{\partial x \partial y}, \; t = \frac{\partial^2 z}{\partial y^2}\right).$$

These equations are special cases of the general elliptic equation

$$\phi(x, y, z, p, q, r, s, t) = 0, \quad \phi_r\phi_t - \phi_s^2 > 0, \quad \phi_r > 0.$$

The literature concerning these equations being very extensive, we shall not attempt to give a complete list of references. The starting point for many more modern researches has been the work of S. Bernstein,† who was the first to prove the analyticity of the solutions of the general equation with ϕ analytic and who was able to obtain a priori bounds for the second and higher derivatives of z in the quasi-linear type in terms of the bounds of $|z|$, $|p|$, $|q|$ and the derivatives of the coefficients. He was also able to prove the existence of the solution of the quasi-linear equation in some very general cases. He assumed that all the data were analytic. However, his papers are very complicated and certain details require modification. On account of the results of J. Horn, L. Lichtenstein, and many others,‡ the restriction of analyticity has been removed. Some very interesting modern work has been done by Leray and Schauder§ in a paper in which they develop a general theory of non-linear functional equations and apply their results to quasi-linear equations.

---

* Presented to the Society, April 11, 1936; received by the editors February 9, 1937.

† Particularly the papers: *Sur la nature analytique des solutions des équations aux dérivées partielles du second ordre*, Mathematische Annalen, vol. 59 (1904), pp. 20–76 and *Sur la généralization du problème de Dirichlet*, Mathematische Annalen, vol. 69 (1910), pp. 82–136.

‡ For an account of some of these results, see the article by L. Lichtenstein on the theory of elliptic partial differential equations in the Encyklopädie der Mathematischen Wissenschaften, vol. II 3^2, pp. 1280–1334.

§ Jean Leray and Jules Schauder, *Topologie et équations fonctionelles*, Annales Scientifiques de l'École Normale Supérieure, vol. 51 (1934), pp. 45–78.

Schauder* has also obtained good a priori bounds for the solutions (and their derivatives) of linear elliptic equations in any number of variables.

In the present paper, an elliptic pair of linear partial differential equations of the form

$$(1)\quad v_x = -(b_2u_x + cu_y + e),\ v_y = au_x + b_1u_y + d,\ 4ac - (b_1 + b_2)^2 \geqq m > 0,$$

is studied. We assume merely that the coefficients are uniformly bounded and measurable. In such a general case, of course, the functions u and v do not possess continuous derivatives but are absolutely continuous in the sense of Tonelli with their derivatives summable with their squares (over interior closed sets). However, certain uniqueness, existence, and compactness theorems are demonstrated and the functions u and v are seen also to satisfy Hölder conditions. These results are immediately used to show that if $z(x, y)$ is a function which minimizes

$$\iint_R f(x, y, z, p, q)dxdy, \qquad (f_{pp}f_{qq} - f_{pq}^2 > 0, f_{pp} > 0)$$

among all functions (for which the integral may be defined) which take on the same boundary values as $z(x, y)$ and if $z(x, y)$ satisfies a Lipschitz condition on R, its first partial derivatives satisfy Hölder conditions. E. Hopf† has already shown that if p and q satisfy Hölder conditions, then the second derivatives satisfy Hölder conditions. In proving this fact, Hopf shows that if the coefficients in (1) satisfy Hölder conditions, the first partial derivatives of u and v satisfy Hölder conditions. The results of this paper concerning the system (1) together with Hopf's result yield very simple proofs of the existence of a solution of the quasi-linear equation in certain cases, a few of which are presented in §6.

The developments of this paper are entirely straightforward, being, for the most part, generalizations of known elementary results analogous to the step from Riemann to Lebesgue integration. The main tools by means of which the Hölder conditions of u and v in (1) are demonstrated are Theorems 1 and 2 of §2 which state roughly that: if $\xi = \xi(x, y)$, $\eta = \eta(x, y)$ is a 1–1 differentiable transformation of a Jordan region $\overline{R}$ into another Jordan region $\overline{\Sigma}$ in which the ratio of the maximum to the minimum magnification is uniformly bounded, then the functions ξ and η and the functions $x(\xi, \eta)$ and $y(\xi, \eta)$ of the inverse satisfy Hölder conditions. Since these Hölder conditions

* J. Schauder, *Über lineare elliptische Differentialgleichungen zweiter Ordnung*, Mathematische Zeitschrift, vol. 38 (1933–34), pp. 257–282.

† E. Hopf, *Zum analytischen Charakter der Lösungen regulärer zweidimensionaler Variationsprobleme*, Mathematische Zeitschrift, vol. 30, pp. 404–413.

are so important in the modern theory of elliptic equations, these two theorems may prove to be an important tool in this field.

We use the following notation: A function $\phi(x, y)$ is said to be of class $C^{(n)}$ if it is continuous together with its partial derivatives of the first n orders. If E is a point set, $\bar{E}$ denotes its closure and E^* its boundary points. If E and F are point sets, the symbol $E \cdot F$ denotes their product, $E+F$ denotes their sum (E and F need not be mutually exclusive), and $E \subset F$ means that E is a subset of F. The symbol $C(P, r)$ denotes the open circular disc with center at P and radius r.

1. **Preliminary definitions and lemmas.** Most of the definitions and lemmas of this section are either found in the literature or are easily deducible from known results. We include the material of this section for completeness.

DEFINITION 1. A function $u(x, y)$ is said to be *strictly absolutely continuous in the sense of Tonelli*† (*A.C.T.*) on a closed rectangle $(a, c; b, d)$ if it is continuous there and

(i) for almost all X, $a \leqq X \leqq b$, $u(X, y)$ is absolutely continuous in y on the interval (c, d), and for almost all Y, $c \leqq Y \leqq d$, $u(x, Y)$ is absolutely continuous in x on the interval (a, b) and

(ii) $V_c^{(y)d}[u(X, y)]$ and $V_a^{(x)b}[u(x, Y)]$ are summable functions of X and Y, respectively, $V_c^{(y)d}[u(X, y)]$, for instance, denoting the variation on (c, d) of $u(X, y)$ considered as a function of y alone. It is clear that these variations are lower semicontinuous in the large lettered variables.

DEFINITION 2. A function $u(x, y)$ is said to be *A.C.T.* on a region R (or $\bar{R}$) if it is continuous there and strictly *A.C.T.* on each closed interior rectangle.

Remark. Evans‡ has shown that every *continuous* "potential function of its generalized derivatives"§ is *A.C.T.* and conversely, so that his theorems concerning the former functions are applicable to the latter. Thus

LEMMA 1.§ *If $u(x, y)$ is A.C.T. on R, then $\partial u/\partial x = u_x$ and $\partial u/\partial y = u_y$ exist almost everywhere in R and are summable over every closed subregion of R.*

LEMMA 2.§ *If $x = x(s, t)$ and $y = y(s, t)$ is a* 1–1 *continuous transformation of class C' of a region R of the (x, y)-plane into a region Σ of the (s, t)-plane, and if $u(x, y)$ is A.C.T. on R, then the function $u[x(s, t), y(s, t)]$ is A.C.T. on Σ and*

† L. Tonelli, *Sulla quadratura delle superficie*, Atti della Reale Accademia dei Lincei, (6), vol. 3 (1926), pp. 357–362, 445–450, 633–638, 714–719.

‡ G. C. Evans, *Complements of potential theory*, II, American Journal of Mathematics, vol. 55 (1933), pp. 29–49.

§ G. C. Evans, *Fundamental points of potential theory*, Rice Institute Pamphlets, vol. 7, No. 4 (1920), pp. 252–329.

$$u_s = u_x x_s + u_y y_s, \qquad u_t = u_x x_t + u_y y_t$$

almost everywhere.

DEFINITION 3. Let D be a region. By the expression *almost all rectangles of D* we mean the totality of rectangles $a \leqq x \leqq b$, $c \leqq y \leqq d$ in D for which a and b do *not* belong to some set of measure zero of values of x, and c and d do not belong to some set of measure zero of values of y. Naturally either or both sets of measure zero may be vacuous.

DEFINITION 4. Let $\phi(D)$ be a function defined on almost all rectangles D of a region R such that whenever $D = D_1 + D_2$, where D_1 and D_2 are admissible rectangles having only an edge in common, we have that $\phi(D) = \phi(D_1) + \phi(D_2)$. We say that $\phi(D)$ is *absolutely continuous* on R if for every $\epsilon > 0$, there exists a $\delta > 0$ such that for every sequence of non-overlapping admissible rectangles $\{D_n\}$ with $\sum[\text{meas } (D_n)] < \delta$, we have $\sum|\phi(D_n)| < \epsilon$.

LEMMA 3.† *Let $f(x, y)$ be summable on R, let D denote a rectangle $(a, c; b, d)$ on which $f(x, y)$ is summable (this being almost all rectangles of R) and define*

$$\phi(D) = \int_c^d [f(b, y) - f(a, y)]dy = \int_{D^*} f dy,$$

$$\psi(D) = \int_a^b [f(x, d) - f(x, c)]dx = -\int_{D^*} f dx.$$

Then a necessary and sufficient condition that $f(x, y)$ be A.C.T. on R is that $f(x, y)$ be continuous and $\phi(D)$ and $\psi(D)$ be absolutely continuous on each subregion Δ for which $\bar{\Delta} \subset R$. When this is true,

$$\phi(D) = \iint_D \frac{\partial f}{\partial x} dxdy, \qquad \psi(D) = \iint_D \frac{\partial f}{\partial y} dxdy$$

for each rectangle in R.

DEFINITION 5. We say that a function $\phi(x, y)$ is of *class L_p* on a region R if $|\phi|^p$ is summable over R.

DEFINITION 6. We say that a function $u(x, y)$ is of *class D_α* on R or $\bar{R}$ if it is *A.C.T.* there and $|u_x|^\alpha$ and $|u_y|^\alpha$ are summable over every closed subregion of R.

LEMMA 4. *Let $\{\phi_n(x, y)\}$, $n = 1, 2, \cdots$, and $\phi(x, y)$ be of class L_p on a rectangle $(a, c; b, d)$ with*

† A summable function $f(x, y)$ for which $\phi(D)$ and $\psi(D)$ are absolutely continuous on each subregion Δ of R for which $\bar{\Delta} \subset R$ is said to be a "potential function of its generalized derivatives" and this lemma is essentially the theorem of Evans mentioned above (*Complements of potential theory*, loc. cit.).

$$\int_a^b \int_c^d |\phi_n(x, y)|^p dxdy \leqq G, \qquad p > 1,$$

where G is independent of n. Let

$$\Phi_n(x, y) = \int_a^x \int_c^y \phi_n(\xi, \eta) d\xi d\eta, \qquad \Phi(x, y) = \int_a^x \int_c^y \phi(\xi, \eta) d\xi d\eta$$

and suppose that the sequence $\{\Phi_n(x, y)\}$ *converges uniformly to* $\Phi(x, y)$. *Suppose also that* $\{A_n(x, y)\}$, $n=1, 2, \cdots$, *and* $A(x, y)$ *are of class* L_q *on* $(a, c; b, d)$, $q=p/(p-1)$, *and that*

$$\lim_{n\to\infty} \int_a^b \int_c^d |A_n - A|^q dxdy = 0.$$

Then

$$\text{(i)} \qquad \int_a^b \int_c^d |\phi|^p dxdy \leqq \liminf_{n\to\infty} \int_a^b \int_c^d |\phi_n|^p dxdy,$$

$$\text{(ii)} \qquad \lim_{n\to\infty} \int_a^b \int_c^d A_n \phi_n dxdy = \int_a^b \int_c^d A\phi dxdy.$$

Proof. The first conclusion is well known.†

To prove (ii), choose $\epsilon>0$. For $n>N_1$, we see from the Hölder inequality that

$$\left|\int_a^b \int_c^d (A_n - A)\phi_n dxdy\right|$$
$$\leqq \left[\int_a^b \int_c^d |A_n - A|^q dxdy\right]^{1/q} \left[\int_a^b \int_c^d |\phi_n|^p dxdy\right]^{1/p} < \frac{\epsilon}{4}.$$

Now, let $\{B_k(x, y)\}$ be a sequence of step functions‡ such that

$$\lim_{k\to\infty} \int_a^b \int_c^d |B_k - A|^q dxdy = 0.$$

Then, for $k>K$ (independent of n),

$$\left|\int_a^b \int_c^d (B_k - A)\phi_n dxdy\right|$$

† For instance, this result may be obtained by the method of proof used in Theorem 7, §1 of the author's paper, *A class of representations of manifolds*, I, American Journal of Mathematics, vol. 55 (1933), p. 693.

‡ To form these, let G_k denote the grating formed by all lines of the form $x=2^{-k}i$, $y=2^{-k}j$. We then define B_k to be a properly chosen constant on the part of each square of G_k which contains a point of the rectangle.

$$\leqq \left[\int_a^b\int_c^d |B_k - A|^q dxdy\right]^{1/q}\left[\int_a^b\int_c^d |\phi_n|^p dxdy\right]^{1/p} < \frac{\epsilon}{4},$$

$$\left|\int_a^b\int_c^d (B_k - A)\phi dxdy\right|$$

$$\leqq \left[\int_a^b\int_c^d |B_k - A|^q dxdy\right]^{1/q}\left[\int_a^b\int_c^d |\phi|^p dxdy\right]^{1/p} < \frac{\epsilon}{4}\cdot$$

Now, let $k_0 > K$. Then, for $n > N_2$,

$$\left|\int_a^b\int_c^d B_{k_0}(\phi_n - \phi)dxdy\right| \leqq \sum_i |B_{k_0,i}| \cdot \left|\iint_{R_i}(\phi_n - \phi)dxdy\right| < \frac{\epsilon}{4},$$

the R_i being the rectangular subregions of R over each of which B_{k_0} is constant and $B_{k_0,i}$ being the value of B_{k_0} on R_i.

Finally, if we let N be the larger of N_1 and N_2, we see that

$$\left|\int_a^b\int_c^d (A_n\phi_n - A\phi)dxdy\right|$$

$$\leqq \left|\int_a^b\int_c^d (A_n - A)\phi_n dxdy\right| + \left|\int_a^b\int_c^d (A - B_{k_0})\phi_n dxdy\right|$$

$$+ \left|\int_a^b\int_c^d B_{k_0}(\phi_n - \phi)dxdy\right| + \left|\int_a^b\int_c^d (B_{k_0} - A)\phi dxdy\right| < \epsilon.$$

This proves the lemma.

LEMMA 5. *Let $\phi(x, y)$ be of class L_p on R, $p \geqq 1$, and let R_h be that subset of points (x_0, y_0) of R such that all points (x, y) with $|x - x_0| < h$, $|y - y_0| < h$ are in R. Let*

$$\phi_h(x, y) = \frac{1}{4h^2}\int_{x-h}^{x+h}\int_{y-h}^{y+h}\phi(\xi, \eta)d\xi d\eta, \qquad h > 0,$$

be defined in R_h. Then $\phi_h(x, y)$ is continuous and of class L_p on R_h and

(i) $$\iint_{R_h} |\phi_h|^p dxdy \leqq \iint_R |\phi|^p dxdy,$$

(ii) $$\lim_{h\to 0}\left[\iint_{R_h} |\phi_h|^p dxdy - \iint_R |\phi|^p dxdy\right]$$

$$= \lim_{h\to 0}\iint_{R_{h_0}} |\phi_h - \phi|^p dxdy = 0, \qquad h_0 > 0.$$

Proof. This lemma is well known.†

Definition 7. Let $u(x, y)$ be defined on R or $\overline{R}$. If u is of class D_2 on R or $\overline{R}$ with $|u_x|^2$ and $|u_y|^2$ summable over the whole of R, we define

$$D(u) = \iint_R (u_x^2 + u_y^2)dxdy.$$

Otherwise we define $D(u) = +\infty$.

Lemma 6. *Let $\{u_n(x, y)\}$ and $\{v_n(x, y)\}$ each be of class D_2 on $\overline{R}$ with $D(u_n)$ and $D(v_n) \leqq G$ independent of n, and suppose that $\{u_n(x, y)\}$ and $\{v_n(x, y)\}$ converge uniformly on $\overline{R}$ to functions $u(x, y)$ and $v(x, y)$, respectively. Let the sequences $\{a_n(x, y)\}$, $\{b_n\}$, $\{c_n\}$, $\{d_n\}$, $\{e_n\}$, $\{f_n\}$, $\{g_n\}$, $\{h_n\}$, $\{k_n\}$, and $\{l_n\}$ be measurable and uniformly bounded and suppose the sequences converge almost everywhere on R to a, b, c, d, e, f, g, h, k, and l respectively. Then*

(i) $u(x, y)$ *and* $v(x, y)$ *are of class* D_2 *on* $\overline{R}$,

(ii) $D(u) \leqq \liminf_{n\to\infty} D(u_n)$, $D(v) \leqq \liminf_{n\to\infty} D(v_n)$,

(iii)
$$\iint_R [(au_x+bu_y+cv_x+dv_y+e)^2+(fu_x+gu_y+hv_x+kv_y+l)^2]dxdy$$
$$\leqq \liminf_{n\to\infty} \iint_R [(a_nu_{nx}+b_nu_{ny}+c_nv_{nx}+d_nv_{ny}+e_n)^2 + (f_nu_{nx}+g_nu_{ny}+h_nv_{nx}+k_nv_{ny}+l_n)^2]dxdy.$$

Proof. Conclusions (i) and (ii) are well known.†

To prove (iii), let M be the uniform bounds for a_n, etc., and let

$$\phi_n = a_nu_{nx} + b_nu_{ny} + c_nv_{nx} + d_nv_{ny} + e_n, \qquad \phi = au_x + bu_y + cv_x + dv_y + e,$$
$$\psi_n = f_nu_{nx} + g_nu_{ny} + h_nv_{nx} + k_nv_{ny} + l_n, \qquad \psi = fu_x + gu_y + hv_x + kv_y + l.$$

Then, for each $h>0$, we see that $\phi_n^{(h)}$, $\psi_n^{(h)}$, $\phi^{(h)}$ and $\psi^{(h)}$ are uniformly bounded on R_h, the proof for $\phi_n^{(h)}$, for instance, employing the Hölder inequality as follows:

$$|\phi_n^{(h)}| \leqq \frac{1}{4h^2}\Bigg[\left|\int_{x-h}^{x+h}\int_{y-h}^{y+h} a_n(\xi, \eta)u_{n\xi}(\xi, \eta)d\xi d\eta\right|$$
$$+\left|\int_{x-h}^{x+h}\int_{y-h}^{y+h} b_nu_{n\eta}d\xi d\eta\right| + \left|\int_{x-h}^{x+h}\int_{y-h}^{y+h} c_nv_{n\xi}d\xi d\eta\right|$$
$$+\left|\int_{x-h}^{x+h}\int_{y-h}^{y+h} d_nv_{n\eta}d\xi d\eta\right| + \left|\int_{x-h}^{x+h}\int_{y-h}^{y+h} e_nd\xi d\eta\right|\Bigg]$$

† See Lemma 1, §1 of the author's paper, loc. cit.

$$\leqq \frac{1}{4h^2}\left\{\left[\int_{x-h}^{x+h}\int_{y-h}^{y+h}|a_n|^2 d\xi d\eta\right]^{1/2}\left[\int_{x-h}^{x+h}\int_{y-h}^{y+h}|u_{n\xi}^2|\,d\xi d\eta\right]^{1/2}\right.$$

$$\left. + * + * + * + 4Mh^2\right\} \leqq \frac{2MG^{1/2}}{h} + M,$$

each $*$ denoting a term similar to the first term. By Lemma 4, ϕ_{nh} and ψ_{nh} converge at each point to ϕ and ψ respectively. Hence,

$$\iint_{R_h}(\phi_h^2 + \psi_h^2)dxdy = \lim_{n\to\infty}\iint_{R_h}(\phi_{nh}^2 + \psi_{nh}^2)dxdy$$

$$\leqq \liminf_{n\to\infty}\iint_R(\phi_n^2 + \psi_n^2)dxdy,$$

using Lemma 5. The statement (iii) follows from Lemma 5 by letting $h\to 0$.

LEMMA 7. *Let f and ϕ be of class D_2 on R, let D be a Jordan subregion of R such that D^* is a rectifiable curve interior to R on which ϕ is of bounded variation. Then $\partial(f, \phi)/\partial(x, y)$ is summable on D and*

$$\int_{D^*} f d\phi = \iint_D \frac{\partial(f, \phi)}{\partial(x, y)} dxdy,$$

the line integral being the ordinary Stieltjes integral.

Proof. That the Jacobian is summable follows from the Schwarz inequality, since $f_x^2 + f_y^2 + \phi_x^2 + \phi_y^2$ is summable over D.

Let f_h and ϕ_k have their usual significance as average functions. If $|h|$, $|k| < \alpha$, f_h and ϕ_k are defined on $\bar{D}$ and moreover f_h is of class C' if $h > 0$ and ϕ_k is of class C' if $k > 0$. Hence, by Green's theorem

$$\int_{D^*} f_h d\phi_k = \iint_D (f_{hx}\phi_{ky} - f_{hy}\phi_{kx})dxdy = -\int_{D^*}\phi_k df_h.$$

Letting k tend to zero, and using Lemmas 4 and 5 and well known theorems on the Stieltjes integral, we see that

$$\int_{D^*} f_h d\phi = -\int_{D^*}\phi df_h = \iint_D (f_{hx}\phi_y - f_{hy}\phi_x)dxdy.$$

The result follows by letting h tend to zero.

2. **Fundamental theorems on transformations.** We state first

DEFINITION 1. We say that $u(x, y)$ satisfies a condition $A[\lambda; M(a, d)]$ on R if it is of class D_2 on R and

$$\iint_{C(P,r)} (u_x^2 + u_y^2)dxdy \leqq M(a, d)\left(\frac{r}{a}\right)^\lambda, \qquad 0 \leqq r \leqq a,$$

$$P = (x, y) \varepsilon R, \qquad \lambda > 0,$$

where $a > 0$, $d > 0$, and $a + d$ is the distance of (x, y) from R^*, $M(a, d)$ depending on a and d and not on (x, y).

DEFINITION 2. We say that $u(x, y)$ satisfies a condition $B[\mu; N(a, d)]$ on R if

$$| u(x_1, y_1) - u(x_2, y_2) | \leqq N(a, d)\left(\frac{r}{a}\right)^\mu, \qquad 0 \leqq r < a,$$

$$r = [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2},$$

provided that every point on the segment joining (x_1, y_1) to (x_2, y_2) is at a distance $\geqq a/2 + d$ from R^*.

LEMMA 1. *Let $u(x, y)$ satisfy a condition $A[\lambda; M(a, d)]$ on R. Then it also satisfies a condition $B[\lambda/2; N(a, d)]$, where*

$$N(a, d) = 8 \cdot 3^{-1/2} \cdot \lambda^{-1} \cdot [M(a/2, d)]^{1/2}.$$

Proof. First assume that $u(x, y)$ is of class C' on R. Let $P_1: (x_1, y_1)$ and $P_2: (x_2, y_2)$ be two points of R which are such that every point on the segment joining them is at a distance $\geqq a/2 + d$ from R^*. Next, choose axes so that P_1 is the origin, and P_2 is the point $(2^{-1/2} \cdot r, 2^{-1/2} \cdot r)$. Then each square of the form

$$0 \leqq x \leqq 2^{-1/2} \cdot rt, \qquad 0 \leqq y \leqq 2^{-1/2} \cdot rt \qquad \text{or}$$

$$2^{-1/2} \cdot r \cdot (1 - t) \leqq x \leqq 2^{-1/2} r, \qquad 2^{-1/2} \cdot r \cdot (1 - t) \leqq y \leqq 2^{-1/2} \cdot r$$

is in a circle of radius $rt/2$ whose center is at a distance $\geqq a/2 + d$ from R^*. Let $\alpha = 2^{-1/2} r$; then

$$\int_0^{\alpha t} \int_0^{\alpha t} (u_x^2 + u_y^2)dxdy \leqq M(a/2, d)(rt/a)^\lambda, \qquad 0 \leqq t \leqq 1,$$

$$\int_{\alpha - \alpha t}^{\alpha} \int_{\alpha - \alpha t}^{\alpha} (u_x^2 + u_y^2)dxdy \leqq M(a/2, d)(rt/a)^\lambda, \qquad 0 \leqq t \leqq 1.$$

Now, for each (x, y) with $0 \leqq x \leqq \alpha$, $0 \leqq y \leqq \alpha$, we have

$$u(\alpha, \alpha) - u(0, 0) = x \int_0^1 u_x(xt, yt)dt + y \int_0^1 u_y(xt, yt)dt$$

$$- (x - \alpha) \int_0^1 u_x[\alpha + t(x - \alpha), \alpha + t(y - \alpha)]dt$$

$$- (y - \alpha) \int_0^1 u_y[\alpha + t(x - \alpha), \alpha + t(y - \alpha)]dt.$$

Integrating both sides of this equation with respect to x and y, we obtain

$$u(\alpha, \alpha) - u(0, 0) = \frac{1}{\alpha^2}\int_0^\alpha \int_0^\alpha \left\{ \int_0^1 xu_x(xt, yt)dt \right\} dxdy + *$$
$$- \frac{1}{\alpha^2}\int_0^\alpha \int_0^\alpha \left\{ \int_0^1 (x - \alpha)u_x[\alpha + t(x - \alpha), \alpha + t(y - \alpha)]dt \right\} dxdy - *$$
$$= \frac{1}{\alpha^2}\int_0^1 \frac{1}{t^3} \left\{ \int_0^{\alpha t} \int_0^{\alpha t} \xi u_\xi(\xi, \eta) d\xi d\eta + * \right.$$
$$\left. - \int_{\alpha-\alpha t}^{\alpha} \int_{\alpha-\alpha t}^{\alpha} (\xi - \alpha) u_\xi(\xi, \eta) d\xi d\eta - * \right\} dt,$$

the last being obtained by suitable changes of variable; the $*$ denotes the term in y or η which is similar to the term preceding it. Using Schwarz's inequality on the interior terms, we obtain

$$| u(\alpha, \alpha) - u(0, 0) | \leqq \int_0^1 \alpha^{-2}t^{-3} \left\{ \left[\int_0^{\alpha t} \int_0^{\alpha t} \xi^2 d\xi d\eta \right]^{1/2} \left[\int_0^{\alpha t} \int_0^{\alpha t} u_\xi^2 d\xi d\eta \right]^{1/2} \right.$$
$$\left. + * + \left[\int_{\alpha-\alpha t}^{\alpha} \int_{\alpha-\alpha t}^{\alpha} (\xi - \alpha)^2 d\xi d\eta \right]^{1/2} \left[\int_{\alpha-\alpha t}^{\alpha} \int_{\alpha-\alpha t}^{\alpha} u_\xi^2 d\xi d\eta \right]^{1/2} + * \right\} dt$$
$$\leqq 4 \cdot 3^{-1/2} [M(a/2, d)]^{1/2} \cdot (r/a)^{\lambda/2} \int_0^1 t^{\lambda/2-1} dt$$
$$= 8\lambda^{-1} 3^{-1/2} [M(a/2, d)]^{1/2} \cdot (r/a)^{\lambda/2}.$$

If $u(x, y)$ is merely of class D_2, $u_h(x, y)$ is of class C' and we obtain the general result by letting h tend to zero.

DEFINITION 3. Let T: $\xi = \xi(x, y)$, $\eta = \eta(x, y)$ be a 1–1 continuous transformation of a closed region $\overline{R}$ into a closed region $\overline{\Sigma}$, which is of class C' in R with $\xi_x\eta_y - \xi_y\eta_x \neq 0$. Let (x_0, y_0) be a point of R, $x = x(\sigma)$, $y = y(\sigma)$ (σ arc length) be a regular curve such that $x(\sigma_0) = x_0$, $y(\sigma_0) = y_0$, $x'(\sigma_0) = \cos\theta$, $y'(\sigma_0) = \sin\theta$. If (ξ_0, η_0) is the point of Σ corresponding to (x_0, y_0) and if ds is the differential of arc length of the curve in Σ corresponding to the above, we define the *magnification of T at (x_0, y_0) in the direction θ by* $|ds/d\theta|$.

Remarks. Clearly this magnification depends only on (x_0, y_0) and θ and not on the curve chosen. It is given by

$$\left| \frac{ds}{d\sigma} \right|^2 = E_0 \cos^2\theta + 2F_0 \sin\theta\cos\theta + G_0 \sin^2\theta, \qquad E_0 = E(x_0, y_0), \text{ etc.}$$

The square of its maximum and minimum (with respect to θ) at P_0 are given by

$$\tfrac{1}{2}\{E_0+G_0+[(E_0+G_0)^2-4(E_0G_0-F_0^2)]^{1/2}\},$$
$$\tfrac{1}{2}\{E_0+G_0-[(E_0+G_0)^2-4(E_0G_0-F_0^2)]^{1/2}\},$$

respectively, so that the ratio of the maximum to the minimum magnification is

$$\mu_0+(\mu_0^2-1)^{1/2}, \qquad \mu_0=\frac{E_0+G_0}{2(E_0G_0-F_0^2)^{1/2}}$$

at P_0. If this ratio is uniformly bounded in R, it is clear that the inverse transformation has the same property. In the foregoing remarks, E, F, and G have their usual differential-geometric significance:

$$E=\xi_x^2+\eta_x^2, \qquad F=\xi_x\xi_y+\eta_x\eta_y, \qquad G=\xi_y^2+\eta_y^2.$$

THEOREM 1. *Let $\overline{R}$ and $\overline{\Sigma}$ be two Jordan regions, a, b, and c be three distinct points of R^*, and α, β, and γ be three distinct points of Σ^*. Let $\{T\}$ be a family of 1–1 continuous transformations of the form*

$$T:\xi=\xi(x,y), \qquad \eta=\eta(x,y)$$

which carry $\overline{R}$ into $\overline{\Sigma}$, which carry a, b, and c into α, β, and γ respectively, and which satisfy the following hypotheses:

(1) *each T is of class C' within R with $\xi_x\eta_y-\xi_y\eta_x\neq 0$, and*

(2) *the ratio of the maximum to the minimum magnification of each transformation at each point (x, y) is $\leqq K$, which is independent of x, y, and T. Then there exist functions $M(a)$, $N(a)$, $P(a)$, and $m(a)$ which depend only on K, the regions $\overline{R}$ and $\overline{\Sigma}$, and the distribution of the points a, b, c, α, β, and γ; and there exists a number $\lambda>0$ which depends only on K such that*

(i) *$M(a)$, $N(a)$, $m(a)>0$ for $a>0$, $\lim_{a\to 0}M(a)=\lim_{a\to 0}N(a)=\lim_{a\to 0}P(a)=\lim_{a\to 0}m(a)=0$,*

(ii) *all points of R or Σ which are at a distance $\geqq\rho>0$ from R^* or Σ^* correspond to points of the other region at a distance $\geqq m(\rho)$ from its boundary, and*

(iii) *the functions $\xi(x,y)$, $\eta(x,y)$, $x(\xi,\eta)$, and $y(\xi,\eta)$ all satisfy conditions of the form $A[2\lambda;\ M(a,d)]$ and $B[\lambda;\ N(a,d)]$ with $M(a,d)=M(a)$, $N(a,d)=N(a)$, and the equicontinuity condition*

$$|\phi(\alpha_1,\beta_1)-\phi(\alpha_2,\beta_2)|\leqq P(a), \qquad [(\alpha_2-\alpha_1)^2+(\beta_2-\beta_1)^2]^{1/2}=a,$$
$$(\alpha,\beta)=(x,y) \text{ or } (\xi,\eta),\ \phi=\xi,\eta,x, \text{ or } y.$$

Proof. Since the ratio of maximum to minimum magnification is $\leqq K$, it follows that

$$\iint_R (\xi_x^2 + \xi_y^2 + \eta_x^2 + \eta_y^2)dxdy \leqq 2Km(\Sigma),$$

$$\iint_\Sigma (x_\xi^2 + x_\eta^2 + y_\xi^2 + y_\eta^2)d\xi d\eta \leqq 2Km(R).$$

From this,† follows the existence of the functions $P(a)$ and $m(a)$ satisfying the desired conditions.

Now, let P_0 belong to R, for example, being at a distance a from R^*. The circle $(x-x_0)^2+(y-y_0)^2 \leqq a^2$ is carried into a Jordan subregion of $\overline{\Sigma}$ which is surely a subset of the circle $(\xi-\xi_0)^2+(\eta-\eta_0)^2 \leqq [P(a)]^2$, (ξ_0, η_0) being the correspondent of (x_0, y_0). Let

$$A(r) = \iint_{C(P_0,r)} J(x, y)dxdy = \int_0^r \int_0^{2\pi} \rho J(x_0 + \rho\cos\theta, y_0 + \rho\sin\theta)d\rho d\theta,$$

$$0 \leqq r \leqq a, \qquad J(x, y) = | \xi_x\eta_y - \xi_y\eta_x |.$$

Then

$$A' = \frac{dA}{dr} = r\int_0^{2\pi} J(x_0 + r\sin\theta, y_0 + r\cos\theta)d\theta \geqq \frac{r}{K}\int_0^{2\pi} (\xi_s^2 + \eta_s^2)d\theta$$

$$= \frac{1}{K}\int_0^{2\pi r} (\xi_s^2 + \eta_s^2)ds \geqq (2K\pi r)^{-1}\left(\int_0^{2\pi r} [\xi_s^2 + \eta_s^2]^{1/2}ds\right)^2$$

$$= (2K\pi r)^{-1}[l(C_r)]^2 \geqq \frac{2}{Kr}\cdot A(r), \qquad 0 < r < a;$$

here, s denotes arc length on the circle $(x-x_0)^2+(y-y_0)^2=r^2$, and C_r is the curve in Σ into which this circle is carried. Thus

$$A'/A \geqq 2/(Kr), \qquad A(a) \leqq \pi[P(a)]^2$$

and hence

$$A(r) \leqq \pi[P(a)]^2(r/a)^{2/K}, \qquad 0 \leqq r \leqq a.$$

Since

$$J(x, y) \geqq (1/2K)\cdot(\xi_x^2 + \xi_y^2 + \eta_x^2 + \eta_y^2),$$

we find that

$$\iint_{C(P_0,r)} (\xi_x^2 + \xi_y^2 + \eta_x^2 + \eta_y^2)dxdy \leqq 2K\pi[P(a)]^2(r/a)^{2/K}.$$

† See the author's paper, *An analytic characterization of surfaces of finite Lebesgue area*, I, American Journal of Mathematics, vol. 57 (1935), Theorem 1, §2, p. 699.

Hence, we see that (iii) is satisfied (remembering Lemma 1) if we choose

$$\lambda = 1/K, \quad M(a) = 2K\pi[P(a)]^2, \quad N(a) = 8\cdot 2^{1/2}\cdot 3^{-1/2}\pi^{1/2}K^{3/2}\cdot P(a/2).$$

DEFINITION 4. We say that a Jordan region $\overline{R}$ and three distinct points a, b, and c of R^* satisfy a condition $D(L, d_0)$ if (1) the distances ab, ac, and bc are all $\geqq d_0>0$, (2) R^* is rectifiable, and (3) if P_1 and P_2 are any two points of R^*, the ratio $\widehat{P_1P_2}\div\overline{P_1P_2}\leqq L$, where $\widehat{P_1P_2}$ is an arc of R^* joining P_1 and P_2 which contains at most one of the points a, b, or c in its interior.

THEOREM 2. *Let the regions $\overline{R}$ and $\overline{\Sigma}$, the points a, b, c, α, β, and γ, and the family $\{T\}$ of transformations satisfy the hypotheses of Theorem 1 and suppose $(\overline{R}; a, b, c)$ and $(\overline{\Sigma}; \alpha, \beta, \gamma)$ satisfy a condition $D(L, d_0)$, $d_0>0$. Then the conclusions of Theorem 1 hold and, in addition, there exists a number M depending only on K, L, d_0 and the areas $m(\Sigma)$ and $m(R)$, and a number $\mu>0$ depending only on K and L such that*

$$\iint_{C(P_0,r)\cdot R}(\xi_x^2+\xi_y^2+\eta_x^2+\eta_y^2)dxdy \leqq Mr^{\mu},$$

$$\iint_{C(P_0,r)\cdot\Sigma}(x_\xi^2+x_\eta^2+y_\xi^2+y_\eta^2)d\xi d\eta \leqq Mr^{\mu},$$

for any point P_0 in the plane.

Proof. We need to prove only the last statement. Let P_0 be a point in the plane and let $0<r\leqq d_0/2$. Then the set $C(P_0, r)\cdot R$ is vacuous or consists of a finite or denumerable number of Jordan regions r_n, the boundary of each of which consists of (1) a finite or denumerable set of arcs of $C^*(P_0, r)\cdot\overline{R}$, (2) a finite or denumerable number of arcs of R^*, and (3) points of R^* which are limit points of all of these arcs. All the points of (2) and (3) are on one of the arcs $\widehat{abc}$, $\widehat{bca}$, or $\widehat{cab}$ of R^*, say $\widehat{abc}$. Clearly r_n^* and $r_{n'}^*$ have at most one point in common if $n\neq n'$. Let $E_{n,r}$ be the set (1) above for each r_n, let $E_r=\sum E_{r,n}$, let σ_n be the region of Σ corresponding to r_n, let $\sigma=\sum\sigma_n$, let $C_{r,n}=\sigma_n^*$, let $C_r=\sum C_{r,n}$, let $\Gamma_{r,n}$ be the totality of arcs corresponding to $E_{r,n}$, and let $\Gamma_r=\sum\Gamma_{r,n}$. Clearly $C_{r,n}\cdot C_{r,n'}$ is at most one point if $n\neq n'$. Let $l(C_r)=\sum l(C_{r,n})$, let $l(\Gamma_{r,n})$ be the sum of the lengths of all the arcs of $\Gamma_{r,n}$ and let $l(\Gamma_r)=\sum l(\Gamma_{r,n})$. It is clear that

$$[l(C_r)]^2 \geqq 4\pi m(\sigma).$$

Consider an r_n, and the closed set $R^*\cdot r_n^*$. Proceeding along the arc $(\widehat{abc})$, there is a first point P_n and a last point Q_n of this set. Then, there is an arc of $E_{r,n}$ joining P_n to Q_n. Hence if Π_n and K_n are the corresponding points

of Σ^*, they are on the arc $\widehat{\alpha\beta\gamma}$ and there is an arc of $\Gamma_{r,n}$ joining them. Now $l(\Pi_n, K_n)$ (arc of $\widehat{\alpha\beta\gamma}$) $\leqq L$ times the length of this arc of $\Gamma_{r,n}$. Hence it is easy to see that

$$l(\Gamma_{r,n}) \geqq \frac{1}{L+1} l(C_{r,n}), \qquad l(\Gamma_r) \geqq \frac{1}{L+1} l(C_r).$$

Any of the above sets may be null and for a set of measure zero of values of r, $l(C_r)$ and $l(\Gamma_r)$ may be infinite.

We may now proceed as in Theorem 1. Let

$$A(r) = \iint_{C(P_0,r)\cdot R} J(x, y)dxdy = \int_0^r \left[\rho \int_{E_\rho} J(x_0 + \rho\cos\theta, y_0 + \rho \sin\theta)d\theta\right] d\rho,$$
$$0 \leqq r \leqq d_0/2, \qquad J(x, y) = |\xi_x\eta_y - \xi_y\eta_x|,$$

(the integral being zero if the field of integration is null). Then

$$\begin{aligned} A'(r) &= r\int_{E_r} J(x_0 + r\cos\theta, y_0 + r\sin\theta)d\theta \geqq \frac{1}{K}\int_{E_r} (\xi_s^2 + \eta_s^2)ds \\ &\geqq (2K\pi r)^{-1}\left[\int_{E_r} (\xi_s^2 + \eta_s^2)^{1/2}ds\right]^2 = (2K\pi r)^{-1}[l(\Gamma_r)]^2 \\ &\geqq (2K\pi r)^{-1}(L+1)^{-2}[l(C_r)]^2 \geqq 2K^{-1}(L+1)^{-2}r^{-1}\cdot A(r). \end{aligned}$$

Also $A(d_0/2) \leqq m(\Sigma)$. Hence as before

$$A(r) \leqq m(\Sigma)\cdot(2r/d_0)^{2/K(L+1)^2},$$

$$\iint_{C(P_0,r)} (\xi_x^2 + \xi_y^2 + \eta_x^2 + \eta_y^2)dxdy \leqq 2Km(\Sigma)(2r/d_0)^{2/K(L+1)^2}.$$

Since $A(r) \leqq m(\Sigma)$ for all values of r, the theorem follows.

LEMMA 2. *Let a, b_1, b_2, c, d, and e be measurable functions defined on a bounded region R with $|a|$, $|b_1|$, $|b_2|$, $|c|$, $|d|$, $|e| \leqq M \geqq 1$ on R. Then there exist sequences $\{a_n\}$, $\{b_{1,n}\}$, $\{b_{2,n}\}$, $\{c_n\}$, $\{d_n\}$, and $\{e_n\}$ which are analytic on $\bar{R}$ and uniformly bounded and which converge almost everywhere on R to a, b_1, b_2, c, d, and e respectively. If $b_1 = b_2$, we may choose $b_{1,n} = b_{2,n}$ for each n. If $b_1 = b_2 = b$ and $ac - b^2 \geqq m > 0$ on R, we may choose the sequences so that there exist numbers $\bar{M}$ and $\bar{m} > 0$ such that*

$$|a_n|, \ |b_n|, \ |c_n|, \ |d_n|, \ |e_n| \leqq \bar{M}, \quad a_nc_n - b_n^2 \geqq \bar{m}.$$

Further, if $ac - b^2 = 1$, the sequences may be chosen so that $a_nc_n - b_n^2 = 1$.

Proof. Let D be a region containing $\bar{R}$ in its interior and define $a = c = 1$,

$b_1=b_2=d=e=0$ in $D-R$. For h sufficiently small, a_h, b_{1h}, b_{2h}, c_h, d_h, and e_h are defined and continuous in a region containing $\bar{R}$ and all are numerically $\leqq M$.

Now suppose $b_1=b_2=b$, $ac-b^2\geqq m>0$. We know that $ac-b^2$ is the product of the maximum by the minimum (for $0\leqq\theta\leqq 2\pi$, (x, y) fixed) of

$$\begin{aligned} f(x, y; \theta) &= a\cos^2\theta + 2b\sin\theta\cos\theta + c\sin^2\theta \\ &= \tfrac{1}{2}\left[(a+c)+(a-c)\cos 2\theta + 2b\sin 2\theta\right]. \end{aligned}$$

Clearly $|f(x, y; \theta)|\leqq 2M$ so that the minimum above $\geqq m/2M$. Thus

$$a_h c_h - b_h^2 \geqq m^2/4M^2 > 0$$

for each $h>0$ and all (x, y) in D_h. The remainder of the proof is obvious.

LEMMA 3. *Let a, b, and c be analytic in a region G which contains the Jordan region $\bar{R}$ in its interior and suppose $|a|$, $|b|$, $|c|\leqq M$, $ac-b^2=1$, $a>0$; let $\bar{\Sigma}$ be another Jordan region. Then there exists a unique* 1–1 *analytic map $\xi=(x, y)$, $\eta=\eta(x, y)$ of $\bar{R}$ on Σ which carries three given distinct points p, q, and r on R^* into three given distinct points π, κ, ρ (arranged in the same order) on Σ^*, and which satisfies*

$$\eta_x = -(b\xi_x + c\xi_y), \qquad \eta_y = a\xi_x + b\xi_y.$$

The Jacobian does not vanish and the ratio of the maximum to the minimum magnification of the transformation is $\leqq M$ at each point.

Proof. Let $\bar{D}$ be a Jordan region contained in G and containing $\bar{R}$ whose boundary is a regular, analytic, simple, closed curve. It is known† that there exists a solution $X(x, y)$ of the equation

$$\frac{\partial}{\partial x}(aX_x + bX_y) + \frac{\partial}{\partial y}(bX_x + cX_y) = 0$$

which takes on the values $X=x$ on D^*, which is analytic on $\bar{D}$, and whose first derivatives do not vanish simultaneously. Clearly there exists an analytic conjugate function $Y(x, y)$ which satisfies the same equation and the relations

$$Y_x = -(bX_x + cX_y), \qquad Y_y = aX_x + bX_y.$$

The equations $X=X(x, y)$, $Y=Y(x, y)$ yield a 1–1 analytic map of ($\bar{D}$ and hence) $\bar{R}$ onto a region $\bar{\Delta}$ which carries p, q, and r into three points π', κ', and ρ' arranged on Δ^* in the same order, and for which $X_xY_y-X_yY_x\neq 0$. If $\xi=\Xi(X, Y)$, $\eta=\mathrm{H}(X, Y)$ is the conformal map of $\bar{\Delta}$ on $\bar{\Sigma}$ which carries π', κ', and ρ' into π, κ, and ρ (respectively), it is easily seen that

† See Lichtenstein, loc. cit.

$$\xi = \xi(x, y) = \Xi[X(x, y), Y(x, y)], \quad \eta = \eta(x, y) = \mathrm{H}[X(x, y), Y(x, y)]$$

is a mapping of the desired type. That this mapping is unique follows from the fact that if $\xi = \xi'(x, y)$, $\eta = \eta'(x, y)$ is another such map, then (ξ', η') are related to (ξ, η) by a conformal transformation with π, κ, and ρ fixed; thus $\xi' \equiv \xi$, $\eta' \equiv \eta$.

LEMMA 4. *Let $\xi = \xi(x, y)$, $\eta = \eta(x, y)$ be a transformation defined on a region R in which the functions ξ and η are of class D_2. Suppose also that*

$$\eta_x = -\xi_y, \qquad \eta_y = \xi_x$$

almost everywhere in R. Then the above map is conformal, i.e., ξ and η are conjugate harmonic functions.

Proof. Let $\bar{D}$ be a rectangle on the boundary of which $\eta(x, y)$ is absolutely continuous, such rectangles being almost all rectangles in R. Then, from Lemma 7, §1, it follows that

$$L(S_D) = \iint_D (EG - F^2)^{1/2} dxdy = \frac{1}{2}\iint_D (\xi_x^2 + \xi_y^2 + \eta_x^2 + \eta_y^2) dxdy$$
$$= \iint_D (\xi_x \eta_y - \xi_y \eta_x) dxdy = \int_{D^*} \xi d\eta,$$

S_D being the surface $\xi = \xi(x, y)$, $\eta = \eta(x, y)$, $(x, y) \varepsilon \bar{D}$, $L(S_D)$ meaning its Lebesgue area. Since the Geöcze area† $G(S)$ of any surface with this boundary curve must be at least as great as $L(S_D)$, and since $L(S) \geqq G(S)$ for every surface, we see that S_D is a surface of minimum area bounded by its boundary curve. Hence $\xi(x, y)$ and $\eta(x, y)$ must be harmonic, since otherwise they could be replaced by the harmonic functions having the same boundary values to form a surface of smaller area bounded by the boundary of S_D.

THEOREM 3. *Let $\bar{R}$ and $\bar{\Sigma}$ be Jordan regions, let p, q, and r be distinct points on R^* and let π, κ, and ρ be distinct points arranged in the same order on Σ^*. Let a, b, and c be bounded, measurable functions defined on R:*

$$|a|, \ |b|, \ |c| \leqq M, \qquad ac - b^2 = 1.$$

Then there exists a 1–1 *continuous transformation $\xi = \xi(x, y)$, $\eta = \eta(x, y)$ of $\bar{R}$ into $\bar{\Sigma}$*

(i) *which carries p, q, and r into π, κ, and ρ respectively,*

(ii) *which is such that $\xi(x, y)$, $\eta(x, y)$, $x(\xi, \eta)$, and $y(\xi, \eta)$ are of class D_2 on $\bar{R}$ and $\bar{\Sigma}$, $x = x(\xi, \eta)$, $y = y(\xi, \eta)$ being the inverse transformation,*

(iii) *in which the conclusions of Theorem* 1 *apply to the functions $\xi(x, y)$,*

† See the author's paper, loc. cit., pp. 696, 698.

$\eta(x, y)$, $x(\xi, \eta)$, and $y(\xi, \eta)$ and in which those of Theorem 2 *also apply if $(\overline{R}; p, q, r)$ and $(\overline{\Sigma}; \pi, \kappa, \rho)$ satisfy a condition $D(L, d_0)$, and*

(iv) *in which the functions $\xi(x, y)$ and $\eta(x, y)$ satisfy*

$$\eta_x = -(b\xi_x + c\xi_y), \qquad \eta_y = a\xi_x + b\xi_y$$

almost everywhere on R.

Proof. Let $\{a_n\}\to a$, $\{b_n\}\to b$, $\{c_n\}\to c$ almost everywhere on R, the a_n, b_n, and c_n being analytic on $\overline{R}$ and satisfying

$$|a_n|,\ |b_n|,\ |c_n| \leqq \overline{M}, \qquad a_n c_n - b_n^2 = 1.$$

Let T_n: $\xi=\xi_n(x, y)$, $\eta=\eta_n(x, y)$ be the unique analytic transformations of $\overline{R}$ into $\overline{\Sigma}$ which carry p, q, and r into π, κ, and ρ respectively and which satisfy

$$\eta_{nx} = -(b_n\xi_{nx} + c_n\xi_{ny}), \qquad \eta_{ny} = a_n\xi_{nx} + b_n\xi_{ny}.$$

Let $x=x_n(\xi, \eta)$, $y=y_n(\xi, \eta)$ denote the inverses. These transformations satisfy the conditions of Theorem 1 and hence the conclusions of Theorem 1 and also of Theorem 2 if $(\overline{R}; p, q, r)$ and $(\overline{\Sigma}; \pi, \kappa, \rho)$ satisfy a condition $D(L, d_0)$; it is easily seen that the ratio of maximum to minimum magnification of T_n is $\leqq\lambda_n+(\lambda_n^2-1)^{1/2}$, $\lambda_n=(a_n+c_n)/2$, and is therefore $\leqq 2\overline{M}$. Hence a subsequence $\{n_k\}$ of the integers $\{n\}$ may be chosen so that $\{\xi_{n_k}\}$, $\{\eta_{n_k}\}$, $\{x_{n_k}\}$, and $\{y_{n_k}\}$ all converge uniformly on $\overline{R}$ and Σ to functions $\xi(x, y)$, $\eta(x, y)$, $x(\xi, \eta)$, and $y(\xi, \eta)$ respectively and $x=x(\xi, \eta)$, $y=y(\xi, \eta)$ is the inverse of T: $\xi=\xi(x, y)$, $\eta=\eta(x, y)$. Clearly T is a 1–1 continuous transformation of $\overline{R}$ into $\overline{\Sigma}$ in which p, q, and r correspond to π, κ, and ρ respectively.

Also, since the ratio of maximum to minimum magnification $\leqq 2\overline{M}$, we have

$$\iint_R (\xi_{nx}^2 + \xi_{ny}^2 + \eta_{nx}^2 + \eta_{ny}^2)\,dxdy \leqq 4\overline{M}m(\Sigma),$$

$$\iint_\Sigma (x_{n\xi}^2 + x_{n\eta}^2 + y_{n\xi}^2 + y_{n\eta}^2)\,d\xi d\eta \leqq 4\overline{M}m(R),$$

so that it follows from Lemma 6, §1 that $\xi(x, y)$, $\eta(x, y)$, $x(\xi, \eta)$, and $y(\xi, \eta)$ are all of class D_2 on $\overline{R}$ and $\overline{\Sigma}$. Using the same lemma, we see that (iii) also holds and that

$$\begin{aligned} 0 &\leqq \iint_R [(\eta_x + b\xi_x + c\xi_y)^2 + (\eta_y - a\xi_x - b\xi_y)^2]\,dxdy \\ &\leqq \liminf_{n\to\infty} \iint_R [(\eta_{nx} + b_n\xi_{nx} + c_n\xi_{ny})^2 + (\eta_{ny} - a_n\xi_{nx} - b_n\xi_{ny})^2]\,dxdy = 0, \end{aligned}$$

so that (iv) is demonstrated.

THEOREM 4. *Let* $(\overline{R}; p, q, r)$, $(\overline{\Sigma}; \pi, \kappa, \rho)$, a, b, *and* c *satisfy the hypotheses of Theorem* 3 *and let* T: $\xi = \xi(x, y)$, $\eta = \eta(x, y)$ *be the transformation derived in that theorem. Then* T *enjoys the following further properties*:

(i) *sets of measure zero and hence measurable sets of* R *and* Σ *correspond, and* $(\xi_x\eta_y - \xi_y\eta_x)$ *and* $(x_\xi y_\eta - x_\eta y_\xi)$ *are defined and* $\neq 0$ *except possibly on a set of measure zero in* R *and* Σ *respectively*;

(ii) *if* $\phi(x, y)$ *and* $\psi(\xi, \eta)$ *are summable on measurable subsets* $D \subseteqq R$ *and* $\Delta \subseteqq \Sigma$, *the functions* $\phi[x(\xi, \eta), y(\xi, \eta)] \cdot (x_\xi y_\eta - x_\eta y_\xi)$ *and* $\psi[\xi(x, y), \eta(x, y)] \cdot (\xi_x\eta_y - \xi_y\eta_x)$ *are summable on* $T(D)$ *and* $T^{-1}(\Delta)$ *respectively, and*

$$\iint_{T(D)} \phi[x(\xi, \eta), y(\xi, \eta)] \cdot (x_\xi y_\eta - x_\eta y_\xi) d\xi d\eta = \iint_D \phi dx dy,$$

$$\iint_{T^{-1}(\Delta)} \psi \cdot (\xi_x \eta_y - \xi_y \eta_x) dx dy = \iint_\Delta \psi d\xi d\eta;$$

(iii) *if* $\phi(x, y)$ *and* $\psi(\xi, \eta)$ *are of class* D_2 *on* R *and* Σ *respectively, then* $\phi[x(\xi, \eta), y(\xi, \eta)]$ *and* $\psi[\xi(x, y), \eta(x, y)]$ *are of class* D_2 *on* Σ *and* R *respectively and*

$$\phi_\xi = \phi_x x_\xi + \phi_y y_\xi, \quad \phi_\eta = \phi_x x_\eta + \phi_y y_\eta, \quad \phi_x = \phi_\xi \xi_x + \phi_\eta \eta_x, \quad \phi_y = \phi_\xi \xi_y + \phi_\eta \eta_y,$$

$$\psi_\xi = \psi_x x_\xi + \psi_y y_\xi, \quad \psi_\eta = \psi_x x_\eta + \psi_y y_\eta, \quad \psi_x = \psi_\xi \xi_x + \psi_\eta \eta_x, \quad \psi_y = \psi_\xi \xi_y + \psi_\eta \eta_y,$$

almost everywhere;

(iv) T *is uniquely determined by* (i), (ii), *and* (iv) *of Theorem* 3.

Proof. Let O be an open set in R, let $O = \sum_{i=1}^{\infty} \overline{R}_i$ where the $\overline{R}_i$ are closed non-overlapping rectangles on each of which $\eta(x, y)$ is absolutely continuous, and let S_i be the *surface* S_i: $\xi = \xi(x, y)$, $\eta = \eta(x, y)$, $(x, y) \varepsilon \overline{R}_i$. Clearly $L(S_i)$ is merely the measure of the closed or open region $\overline{\Sigma}_i$ or Σ_i into which $\overline{R}_i$ or R_i is carried, Σ_i^* being a rectifiable curve. From Lemma 7, §1, it follows that

$$L(S_i) = m(\Sigma_i) = \int_{\Sigma_i^*} \xi d\eta = \int_{R_i^*} \xi d\eta = \iint_{R_i} (\xi_x \eta_y - \xi_y \eta_x) dx dy.$$

Thus $\xi_x\eta_y - \xi_y\eta_x \geqq 0$, and if $\Omega = T(O)$, then

$$m(\Omega) = \iint_O (\xi_x \eta_y - \xi_y \eta_x) dx dy.$$

It follows very easily that a measurable set E in R is carried into one in Σ and

$$m[T(E)] = \iint_E (\xi_x \eta_y - \xi_y \eta_x) dx dy.$$

The same proof establishes the fact that a measurable set Δ in Σ is carried into a measurable set D in R and that

$$m(D) = \iint_\Delta (x_\xi y_\eta - x_\eta y_\xi) d\xi d\eta.$$

Hence the rest of (i) follows easily and we have proved (ii) for the case $\phi = \psi = 1$.

Suppose ϕ, for instance, is bounded. Let $\{\phi_n(x, y)\}$ be a sequence of step functions which are uniformly bounded and converge to ϕ almost everywhere. It is clear that the functions $\phi_n[x(\xi, \eta), y(\xi, \eta)] \cdot (x_\xi y_\eta - x_\eta y_\xi)$ and $\phi[x(\xi, \eta), y(\xi, \eta)] \cdot (x_\xi y_\eta - x_\eta y_\xi)$ are all summable and dominated by a summable function, and, for each n, it is clear from the above that

$$\iint_D \phi_n(x, y) dx dy = \iint_{T(D)} \phi_n[x(\xi, \eta), y(\xi, \eta)](x_\xi y_\eta - x_\eta y_\xi) d\xi d\eta.$$

The result (ii) for ϕ follows by a passage to the limit, since $\phi_n[x(\xi, \eta), y(\xi, \eta)]$ converges almost everywhere on Σ to $\phi[x(\xi, \eta), y(\xi, \eta)]$ and the latter function is certainly measurable. If ϕ is merely summable, let $\phi_1 = \phi$ where $\phi \geqq 0$, $\phi_1 = 0$ elsewhere, $\phi_2 = -\phi$ where $\phi < 0$, and $\phi_2 = 0$ elsewhere, let $\phi_{1,N} = \phi_1$ if $\phi_1 \leqq N$, $\phi_{1,N} = N$ elsewhere, $\phi_{2,N} = \phi_2$ if $\phi_2 \leqq N$, $\phi_{2,N} = 0$ elsewhere. Then it follows that the functions

$$\phi_{1,N}(x_\xi y_\eta - x_\eta y_\xi), \qquad \phi_{2,N}(x_\xi y_\eta - x_\eta y_\xi)$$

form monotone non-decreasing sequences of non-negative summable functions converging almost everywhere to $\phi_1 \cdot (x_\xi y_\eta - x_\eta y_\xi)$ and $\phi_2 \cdot (x_\xi y_\eta - x_\eta y_\xi)$ respectively, their integrals remaining bounded. Hence (ii) follows.

To prove (iii), let $\phi(x, y)$ be of class D_2 on R, let G be a subregion of R for which $\overline{G} \subset R$, and let Γ be the corresponding subregion of $\Sigma (\overline{\Gamma} \subset \Sigma$ clearly). Now, let Δ be a rectangle of Γ along which $x(\xi, \eta)$ and $y(\xi, \eta)$ are absolutely continuous. Then, using Lemma 7, §1, we see that

$$\int_{\Delta^*} \phi[x(\xi, \eta), y(\xi, \eta)] d\eta = \int_{D^*} \phi(x, y) d\eta = \iint_D (\phi_x \eta_y - \phi_y \eta_x) dx dy,$$

$$-\int_{\Delta^*} \phi[x(\xi, \eta), y(\xi, \eta)] d\xi = -\int_{D^*} \phi(x, y) d\xi = -\iint_D (\phi_x \xi_y - \phi_y \xi_x) dx dy.$$

Clearly these relations follow for any rectangle Δ of G so that ϕ is $A.C.T.$ as a function of ξ and η by Lemma 3, §1. Using the same lemma, it follows that

$$\iint_\Omega \frac{\partial \phi}{\partial \xi} d\xi d\eta = \iint_O \frac{\partial \phi}{\partial \xi} \cdot (\xi_x \eta_y - \xi_y \eta_x) dx dy = \iint_O (\phi_x \eta_y - \phi_y \eta_x) dx dy,$$

$$\iint_{\Omega} \frac{\partial \phi}{\partial \eta} d\xi d\eta = \iint_{O} \frac{\partial \phi}{\partial \eta} \cdot (\xi_x \eta_y - \xi_y \eta_x) dx dy = \iint_{O} (-\phi_x \xi_y + \phi_y \xi_x) dx dy,$$

for every open set Ω in Σ with $\overline{\Omega} \subset \Sigma$. Hence, almost everywhere in R and Σ,

$$\phi_x \eta_y - \phi_y \eta_x = \phi_\xi \cdot (\xi_x \eta_y - \xi_y \eta_x),$$
$$-\phi_x \xi_y + \phi_y \xi_x = \phi_\eta \cdot (\xi_x \eta_y - \xi_y \eta_x).$$

Since $\xi_x \eta_y - \xi_y \eta_x \neq 0$ except on a set of measure zero, it follows that

$$\phi_x = \phi_\xi \xi_x + \phi_\eta \eta_x, \qquad \phi_y = \phi_\xi \xi_y + \phi_\eta \eta_y,$$

almost everywhere on R and Σ. Setting $\phi = x$ and y in turn, we obtain the relations

$$x_\xi \xi_x + x_\eta \eta_x = 1, \quad x_\xi \xi_y + x_\eta \eta_y = 0, \quad y_\xi \xi_x + y_\eta \eta_x = 0, \quad y_\xi \xi_y + y_\eta \eta_y = 1$$

almost everywhere, from which it follows that

$$\phi_\xi = \phi_x x_\xi + \phi_y y_\xi, \qquad \phi_\eta = \phi_x x_\eta + \phi_y y_\eta$$

almost everywhere on R and Σ; the relations for ψ are proved similarly. It follows that

$$(\xi_x \eta_y - \xi_y \eta_x) \cdot (x_\xi y_\eta - x_\eta y_\xi) = 1$$

almost everywhere on R and Σ.

To show that ϕ is of class D_2 in (ξ, η) we see that

$$(\phi_\xi^2 + \phi_\eta^2) \cdot (\xi_x \eta_y - \xi_y \eta_x) = a\phi_x^2 + 2b\phi_x \phi_y + c\phi_y^2 \leqq 2\overline{M}(\phi_x^2 + \phi_y^2),$$

using the relations above and the relations $\eta_x = -(b\xi_x + c\xi_y)$, $\eta_y = a\xi_x + b\xi_y$, $ac - b^2 = 1$. The fact for ψ as a function of (x, y) can be obtained from the above and the fact that

$$a\psi_x^2 + 2b\psi_x \psi_y + c\psi_y^2 \geqq \frac{1}{2\overline{M}} (\psi_x^2 + \psi_y^2).$$

To show (iv), let $\xi = \xi'(x, y)$, $\eta = \eta'(x, y)$ be another map of $\overline{R}$ into $\overline{\Sigma}$ satisfying (i), (ii), and (iv) of Theorem 3. Then ξ' and η' are of class D_2 in ξ and η and all the rules for differentiation apply. We see that $\eta'_\eta = \xi'_\xi$ and $\eta'_\xi = -\xi'_\eta$ almost everywhere on Σ so the map $\xi' = \xi'(\xi, \eta)$, $\eta' = \eta'(\xi, \eta)$ is conformal by Lemma 4. Hence $\xi' \equiv \xi$, $\eta' \equiv \eta$.

3. **A special elliptic system of partial differential equations of the first order.** We prove first

THEOREM 1. *Let $D(x, y)$ and $E(x, y)$ be of class L_2 on $\Sigma_a = C(0, 0; a)$ such that*

$$\iint_{C(P,\rho)\cdot\Sigma_a} D^2 dxdy \leqq M\rho^\lambda, \qquad \iint_{C(P,\rho)\cdot\Sigma_a} E^2 dxdy \leqq M\rho^\lambda, \qquad \lambda > 0.$$

Then the function

$$U(x, y) = \iint_{\Sigma_a} \frac{(\xi - x)D(\xi, \eta) + (\eta - y)E(\xi, \eta)}{(\xi - x)^2 + (\eta - y)^2} d\xi d\eta$$

is defined and continuous over the whole space, and satisfies a uniform Hölder condition of the form

$$|U(x_1, y_1) - U(x_2, y_2)| \leqq Nr^{\lambda/2}, \quad N = 12(2\pi M)^{1/2}[(2 - \lambda)^{-1} + \lambda^{-1}(\lambda + 1)],$$
$$r = [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2};$$

and

$$|U(x, y)| \leqq 2(1 + \lambda^{-1})(2\pi M)^{1/2}\cdot(2a)^{\lambda/2}, \qquad (x, y)\varepsilon\overline{\Sigma}_a.$$

Proof. Define $D(x, y) = 0$ for $x^2 + y^2 \geqq a^2$ and let

$$U_1(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{(\xi - x)\cdot D}{(\xi - x)^2 + (\eta - y)^2} d\xi d\eta$$
$$= \iint_{\Sigma_a} \frac{(\xi - x)\cdot D}{(\xi - x)^2 + (\eta - y)^2} d\xi d\eta.$$

Let

$$h(x, y; r) = r^{1/2}\int_0^{2\pi} |D(x + r\cos\theta, y + r\sin\theta)\cdot\cos\theta|\, d\theta.$$

Then

$$h^2(x, y; r) = r\left[\int_0^{2\pi} |D\cos\theta|\, d\theta\right]^2.$$

By the Hölder inequality $h^2(x, y; r)$ is summable on any interval $0 \leqq r \leqq h$, since

$$2\pi\iint_{C(x,y;h)} D^2 d\xi d\eta = 2\pi\int_0^h\int_0^{2\pi} rD^2 drd\theta \geqq 2\pi\int_0^h r\left[\int_0^{2\pi} D^2\cos^2\theta d\theta\right] dr$$
$$\geqq \int_0^h h^2(x, y; r)dr.$$

Thus

$$\int_0^h h^2(x, y; r)dr \leqq 2\pi M h^\lambda.$$

If we let $H(x, y; r) = \int_0^r h(x, y; \rho)d\rho$, we find by the Hölder inequality that

$$H(x, y; r) \leqq (2\pi M)^{1/2} r^{(1+\lambda)/2}.$$

If N is so large that Σ_a is in the circle $C(x, y; N)$, then $H(x, y; r) \leqq (2\pi M N a^\lambda)^{1/2}$ for $r \geqq N$. Hence

$$\begin{aligned}\left| \int_\epsilon^\rho \int_0^{2\pi} \frac{(\xi - x)D}{(\xi - x)^2 + (\eta - y)^2} d\xi d\eta \right| &\leqq \int_\epsilon^\rho r^{-1/2} h(x, y; r)dr \\ &= \rho^{-1/2}H(x, y; \rho) - \epsilon^{-1/2}H(x, y; \epsilon) + \frac{1}{2}\int_\epsilon^\rho r^{-3/2}H(x, y; r)dr \\ &\leqq (1 + \lambda^{-1})(2\pi M)^{1/2}\rho^{\lambda/2}\end{aligned}$$

for every $\epsilon > 0$, and every $\rho > 0$. Hence $U_1(x, y)$ is defined.

Now, let P_1 and P_2 be any two points and let $\overline{P}$ be the midpoint of the segment P_1P_2. Let $\rho = \overline{P_1P_2}/2$ and let a circle $\sigma(k\rho)$ of radius $k\rho$, $k > 1$, be described about $\overline{P}$ as center. Let α be the inclination of the segment P_1P_2. Then

$$\begin{aligned}U_1(P_2) - U_1(P_1) = &\iint_{\sigma(k\rho)} \frac{(\xi - x_2)D(\xi, \eta)}{(\xi - x_2)^2 + (\eta - y_2)^2} d\xi d\eta \\ &- \iint_{\sigma(k\rho)} \frac{(\xi - x_1)D(\xi, \eta)}{(\xi - x_1)^2 + (\eta - y_1)^2} d\xi d\eta \\ &+ \int_{-\rho}^{\rho}\left[\iint_{W-\sigma(k\rho)} \frac{\cos(2\theta - \alpha)D(\xi, \eta)}{r^2} d\xi d\eta\right] ds,\end{aligned}$$

where

$$r^2 = (\xi - \bar{x} - s\cdot\cos\ \alpha)^2 + (\eta - \bar{y} - s\cdot\sin\ \alpha)^2, \quad \theta = \tan^{-1}\frac{\eta - \bar{y} - s\cdot\sin\alpha}{\xi - \bar{x} - s\cdot\cos\alpha}.$$

Thus

$$\begin{aligned}|U_1(P_1) - U_1(P_2)| \leqq &\int_0^{(k+1)\rho} r^{-1/2}h(x_1, y_1; r)dr + \int_0^{(k+1)\rho} r^{-1/2}h(x_2, y_2; r)dr \\ &+ \int_{-\rho}^{\rho}\left[\int_{(k-1)\rho}^{\infty} r^{-3/2}h[x(s), y(s); r]dr\right]ds \\ \leqq &\ 2(1 + \lambda^{-1})(2\pi M)^{1/2}(k + 1)^{\lambda/2}\rho^{\lambda/2} \\ &+ \int_{-\rho}^{\rho}\frac{3}{2}\int_{(k-1)\rho}^{\infty} r^{-5/2}H[x(s), y(s); r]drds \\ \leqq &\ 6(2\pi M)^{1/2}(1 + \lambda^{-1} + [2 - \lambda]^{-1})\rho^{\lambda/2} \\ &\quad (k = 2,\ x(s) = \bar{x} + s\cos\alpha,\ y(s) = \bar{y} + s\sin\alpha).\end{aligned}$$

The result follows by using the similar result for $U-U_1$.

LEMMA 1. *Let $D(x, y)$ and $E(x, y)$ be of class L_2 on Σ_a with*

$$\iint_{C(P,\rho)\cdot\Sigma_a} D^2 dxdy \leqq M\rho^\lambda, \qquad \iint_{C(P,\rho)\cdot\Sigma_a} E^2 dxdy \leqq M\rho^\lambda, \qquad \lambda > 0, \qquad \rho > 0.$$

Then there exist sequences $\{D_n(x, y)\}$ and $\{E_n(x, y)\}$ of functions analytic on $\overline{\Sigma}_a$ satisfying the above condition with the same M and λ, and such that

$$\lim_{n\to\infty} \iint_{\Sigma_a} [(D_n - D)^2 + (E_n - E)^2] dxdy = 0.$$

Proof. Define $D(x, y) = E(x, y) = 0$ for $x^2+y^2 \geqq a^2$, and let D_h and E_h be the usual average functions. Then D_h and E_h are continuous on Σ_a and

$$\lim_{h\to\infty} \iint_{\Sigma_a} [(D_h - D)^2 + (E_h - E)^2] dxdy = 0,$$

by Lemma 5, §1. Moreover for each $h>0$

$$\begin{aligned} \iint_{C(x_0,y_0;\rho)} D_h^2\, dxdy &= \iint_{C(x_0,y_0;\rho)} \frac{1}{16h^4}\left[\int_{x-h}^{x+h}\int_{y-h}^{y+h} D d\xi d\eta\right]^2 dxdy \\ &\leqq \frac{1}{4h^2}\iint_{C(x_0,y_0;\rho)} \left[\int_{x-h}^{x+h}\int_{y-h}^{y+h} D^2 d\xi d\eta\right] dxdy \\ &= \frac{1}{4h^2}\int_{-h}^{h}\int_{-h}^{h}\left[\iint_{C(x_0+\xi,y_0+\eta;\rho)} D^2 dxdy\right] d\xi d\eta \end{aligned}$$

the same being true for E. The lemma follows easily from this.

LEMMA 2. *Suppose that D_n, E_n, D, and E are all of class L_2 on Σ_a, satisfy the condition of Lemma 1, M and λ being independent of n, and the condition*

$$\lim_{n\to\infty} \iint_{\Sigma_a} [(D_n - D)^2 + (E_n - E)^2] dxdy = 0.$$

Then the functions

$$U_n(x, y) = \frac{1}{2\pi}\iint_{\Sigma_a} \frac{(\xi - x)D_n + (\eta - y)E_n}{(\xi - x)^2 + (\eta - y)^2} d\xi d\eta$$

converge uniformly on any bounded plane region to the function

$$U(x, y) = \frac{1}{2\pi}\iint_{\Sigma_a} \frac{(\xi - x)D + (\eta - y)E}{(\xi - x)^2 + (\eta - y)^2} d\xi d\eta.$$

Proof. Since the $U_n(x, y)$ are equicontinuous on any bounded set (in fact the whole plane) it is sufficient to prove the convergence at each point $P_0:(x_0, y_0)$.

Choose $\epsilon>0$, and choose ρ_0 so small that

$$\left| \frac{1}{2\pi} \iint_{C(P_0,\rho_0)} \frac{(\xi - x_0)D_n + (\eta - y_0)E_n}{(\xi - x_0)^2 + (\eta - y_0)^2} d\xi d\eta \right| + \left| \frac{1}{2\pi} \iint_{C(P_0,\rho_0)} \frac{(\xi - x_0)D + (\eta - y_0)E}{(\xi - x_0)^2 + (\eta - y_0)^2} d\xi d\eta \right| < \frac{\epsilon}{2} \cdot$$

Then there exists an N_0 such that for $n>N_0$,

$$\begin{aligned} &\left| \frac{1}{2\pi} \iint_{\Sigma_a - \Sigma_a \cdot C(P_0,\rho_0)} \frac{(\xi - x_0)(D_n - D) + (\eta - y_0)(E_n - E)}{(\xi - x_0)^2 + (\eta - y_0)^2} d\xi d\eta \right| \\ &\qquad \leqq \frac{1}{2\pi\rho_0} \iint_{\Sigma_a} [(D_n - D)^2 + (E_n - E)^2]^{1/2} d\xi d\eta \\ &\qquad \leqq (2\rho_0)^{-1} \cdot \pi^{-1/2} \cdot a \cdot \left[\iint_{\Sigma_a} [(D_n - D)^2 + (E_n - E)^2] d\xi d\eta \right]^{1/2} < \frac{\epsilon}{2} \cdot \end{aligned}$$

This proves the lemma.

Definition 1. Let D and E be functions of class L_2 on Σ_a. Then if u and v are functions of class D_2 on Σ_a, we define

$$K(u, v) = \iint_{\Sigma_a} [v_x(u_x + D) + v_y(u_y + E)] dxdy,$$

$$J(u) = \iint_{\Sigma_a} [(u_x + D)^2 + (u_y + E)^2] dxdy.$$

If the functions D and E have subscripts, we shall denote the integrals formed by using the new functions by putting the same subscripts on J and K.

Remarks. It is clear that $J(u+\zeta) = J(u) + 2K(u, \zeta) + D(\zeta)$. Also

$$J(u) \leqq 2D(u) + 2 \iint_{\Sigma_a} (D^2 + E^2) dxdy,$$

$$D(u) \leqq 2J(u) + 2 \iint_{\Sigma_a} (D^2 + E^2) dxdy.$$

Accordingly, if H is the harmonic function which takes on the same boundary values as u, we see that

$$J(u) \leqq J(H) \leqq 2D(H) + 2\iint_{\Sigma_a} (D^2 + E^2)dxdy;$$

$$D(u) \leqq 4D(H) + 6\iint_{\Sigma_a} (D^2 + E^2)dxdy.$$

THEOREM 2. *Let D and E satisfy the conditions of Lemma 1, and let u^* be a continuous function defined on Σ_a^* for which $D(H^*)$ is finite, H^* being the harmonic function which takes on the given boundary values. Then there exists a unique function u of class D_2 on $\overline{\Sigma}_a$ which takes on these boundary values and minimizes $J(u)$ among all such functions. The function $u(x, y)$ is given by*

$$u(x, y) = H_a(x, y) + U_a(x, y),$$

$$U_a(x, y) = \frac{1}{2\pi}\iint_{\Sigma_a} \frac{(\xi - x)D + (\eta - y)E}{(\xi - x)^2 + (\eta - y)^2}\, d\xi d\eta,$$

$H_a(x, y)$ being the harmonic function which takes on the boundary values $u^* - U_a$.

Proof. Let $\{D_n\}$ and $\{E_n\}$ be sequences of functions, analytic on $\overline{\Sigma}_a$ and satisfying the conditions of Lemma 2 with respect to D and E. Let $U(x, y)$ be any function of class D_2 on $\overline{\Sigma}_a$ taking on the given boundary values and let $u_n(x, y)$ be the unique solution, for each n, of

$$\Delta u_n + D_{nx} + E_{ny} = 0 \qquad (\Delta U = U_{xx} + U_{yy})$$

which takes on the given boundary values; each u_n is the minimizing function for $J_n(u)$. Then $J_n(U) \to J(U)$, $J_n(U) \geqq J_n(u_n)$. Thus $J(U) \geqq \limsup_{n\to\infty} J_n(u_n)$.

On the other hand, the functions $u_n(x, y)$ are given by

$$u_n(x, y) = H_{a,n}(x, y) + U_{a,n}(x, y),$$

$$U_{a,n}(x, y) = \frac{1}{2\pi}\iint_{\Sigma_a} \frac{(\xi - x)D_n + (\eta - y)E_n}{(\xi - x)^2 + (\eta - y)^2}\, d\xi d\eta,$$

where $U_{a,n}(x, y)$ tends uniformly to $U_a(x, y)$ so that $H_{a,n}(x, y)$ tends uniformly to $H_a(x, y)$ and hence $u_n(x, y)$ tends uniformly to the above $u(x, y)$. Since $D(u_n)$ is uniformly bounded, $u(x, y)$ is of class D_2 and $J(u) \leqq \liminf_{n\to\infty} J_n(u_n)$ by Lemma 6, §1. Hence $u(x, y)$ minimizes $J(u)$.

Let ζ be of class D_2 on $\overline{\Sigma}_a$ and zero on Σ_a^*. Then

$$J(u + \lambda\zeta) = J(u) + 2\lambda K(u, \zeta) + \lambda^2 D(\zeta).$$

Since $u(x, y)$ gives J a minimum, the middle term must vanish. Thus

$$J(u + \zeta) = J(u) + D(\zeta) > J(u)$$

unless $\zeta \equiv 0$.

Remarks. If we let $u_{0,a}(x, y)$ be the minimizing function which is zero on Σ_a^*, we see that it is of class D_2. If D_n and E_n are as in Lemma 2, it is clear that $u_{0,a,n}$ converges uniformly to $u_{0,a}$ and $D(u_{0,a,n})$ is uniformly bounded.

THEOREM 3. *Let D and E satisfy the conditions of Lemma* 1 *and let u be the minimizing function for $J(u)$ which takes on given continuous boundary values for which $J(u)$ is finite. Then there exists a unique function $v(x, y)$ which is of class D_2 on Σ_a, vanishes at the origin, and satisfies*

$$v_x = -(u_y + E), \qquad v_y = u_x + D$$

almost everywhere. Any other function $\overline{V}(x, y)$ which is of class D_2 and satisfies these relations, differs from v by a constant. The function $v(x, y)$ is given by

$$v(x, y) = K_a(x, y) + V_a(x, y),$$

$$V_a(x, y) = \frac{1}{2\pi}\iint_{\Sigma_a} \frac{(\xi - x)E - (\eta - y)D}{(\xi - x)^2 + (\eta - y)^2}\, d\xi d\eta,$$

where $K_a(x, y)$ is the conjugate of the $H_a(x, y)$ of Theorem 2.

Proof. Choose $\{D_n\}$ and $\{E_n\}$ as in Lemma 2, let $u_n(x, y)$ be the minimizing function for $J_n(U)$ with the given boundary values, and let $v_n(x, y)$ satisfy

$$v_{nx} = -(u_{ny} + E_n), \qquad v_{ny} = u_{nx} + D_n.$$

Then

$$v_n(x, y) = K_{a,n}(x, y) + V_{a,n}(x, y),$$

$$V_{a,n}(x, y) = \frac{1}{2\pi}\iint_{\Sigma_a} \frac{(\xi - x)E_n - (\eta - y)D_n}{(\xi - x)^2 + (\eta - y)^2}\, d\xi d\eta.$$

It is seen by well known methods of differentiating the functions $U_{a,n}$ and $V_{a,n}$ that

$$V_{a,n,x} = -(U_{a,n,y} + E_n), \qquad V_{a,n,y} = U_{a,n,x} + D_n$$

on Σ_a. Thus it is clear that $K_{a,n}(x, y)$ is the conjugate of $H_{a,n}(x, y)$ for each n. Since we saw that $H_{a,n}(x, y)$ converged uniformly to $H_a(x, y)$ in Theorem 2, and since $V_{a,n}(x, y)$ obviously tends uniformly to $V_a(x, y)$, it is clear that $K_{a,n}(x, y)$ tends uniformly to $K_a(x, y)$ and $v_n(x, y)$ tends uniformly to the above $v(x, y)$ on each closed subregion of Σ_a. Also $D(v_n) = J(u_n)$ and hence $D(v_n)$ is uniformly bounded.

Now, let $\overline{R}$ be a closed subregion of Σ_a. Then $v(x, y)$ is of class D_2 on $\overline{R}$ with $D(v) \leq \liminf_{n\to\infty} D(v_n)$ which are uniformly bounded. Hence, by Lemma 6, §1,

$$0 \leqq \iint_R [(v_x + u_y + E)^2 + (v_y - u_x - D)^2]dxdy$$

$$\leqq \liminf_{n\to\infty} \iint_R [(v_{nx} + u_{ny} + E_n)^2 + (v_{ny} - u_{nx} - D_n)^2]dxdy = 0.$$

Hence $v(x, y)$ satisfies the desired relations almost everywhere. Now if $\overline{V}$ is of class D_2 and satisfies the same relations almost everywhere, $\overline{V}_x - v_x = \overline{V}_y - v_y = 0$ almost everywhere so that $\overline{V} - v$ is a constant.

THEOREM 4. *If D and E satisfy the conditions of Lemma* 1 *and $u(x, y)$ and $v(x, y)$ are of class D_2 on Σ_a and satisfy*

$$v_x = -(u_y + E), \qquad v_y = u_x + D$$

almost everywhere, then

$$u(x, y) = H_a(x, y) + U_a(x, y), \qquad v(x, y) = K_a(x, y) + V_a(x, y),$$

where U_a and V_a have their previous significance and H_a and K_a are conjugate harmonic functions.

Proof. Let $b<a$ so that v is absolutely continuous on Σ_b^*, this being true for all values of $b<a$ excepting those in a certain set of measure zero (using Lemma 2, §1). Then if ζ is of class D_2 on $\overline{\Sigma}_b$, we have

$$\iint_{\Sigma_b} (\zeta_x v_y - \zeta_y v_x)dxdy = \int_{\Sigma_b^*} \zeta dv.$$

Hence if ζ is also zero on Σ_b^*, we see that

$$\iint_{\Sigma_b} [(u_x + D)\zeta_x + (u_y + E)\zeta_y]dxdy = \iint_{\Sigma_b} (\zeta_x v_y - \zeta_y v_x)dxdy = 0.$$

Thus u is the minimizing function for $J(u)$ on Σ_b with $J(u)$ finite, and $v(x, y)$ is its "conjugate" as in Theorem 3. Hence

$$\begin{aligned} u(x, y) &= H_b'(x, y) + \frac{1}{2\pi}\iint_{\Sigma_b} \frac{(\xi - x)D + (\eta - y)E}{(\xi - x)^2 + (\eta - y)^2} d\xi d\eta \\ &= H_b(x, y) + U_a(x, y), \\ v(x, y) &= K_b'(x, y) + \frac{1}{2\pi}\iint_{\Sigma_b} \frac{(\xi - x)E - (\eta - y)D}{(\xi - x)^2 + (\eta - y)^2} d\xi d\eta \\ &= K_b(x, y) + V_a(x, y). \end{aligned}$$

Clearly H_b and K_b are independent of b, and, since H_b' and K_b' are conjugate harmonic functions, it is easily seen that H_b and K_b are.

4. A more general linear elliptic pair of partial differential equations. In this section, we consider the pair of equations

$$(1) \qquad v_x = -(b_2 u_x + c u_y + e), \qquad v_y = a u_x + b_1 u_y + d,$$

where we assume that a, b_1, b_2, c, d, and e are measurable on a bounded Jordan region R with

$$(2) \quad |a|, |b_1|, |b_2|, |c|, |d|, |e| \leqq M, \ 4ac - (b_1 + b_2)^2 \geqq m > 0, \ a > 0.$$

THEOREM 1. *Let R be a bounded Jordan region and H^* a function continuous on $\overline{R}$ and harmonic on R for which $D(H^*)$ is finite, and suppose $d=e=0$ on R. Then there exists a unique pair of functions $u(x, y)$ and $v(x, y)$*

(i) *which are of class D_2 on $\overline{R}$ and R respectively,*

(ii) *which satisfy* (1) *almost everywhere on R, and*

(iii) *for which $u(x, y) = H^*(x, y)$ on R^* and $v(x_0, y_0) = 0$. These functions also satisfy*

(iv) *conditions $A[2\lambda, M'(\alpha, \delta)]$ and $B[\lambda, N'(\alpha, \delta)]$ on R where λ depends only on M and m, and $M'(\alpha, \delta)$, $N'(\alpha, \delta)$ depend only on M, m, $\overline{R}$, and the maximum of $|H^*|$ on $\overline{R}$, and the maximum of $|u|$ depends only on the bound for $|H^*|$ on R^*.*

Proof. Approximate to a, b_1, b_2, c, d, and e by sequences $\{a_n\}$, $\{b_{1,n}\}$, $\{b_{2,n}\}$, $\{c_n\}$, of functions analytic on $\overline{R}$ for which

$$|a_n|, |b_{1,n}|, |b_{2,n}|, |c_n| \leqq \overline{M}, \qquad 4a_n c_n - (b_{1,n} + b_{2,n})^2 \geqq \overline{m} > 0,$$

and approximate to H^* by harmonic functions H_n^* for which $D(H_n^*) \leqq G$ and such that the solution u_n of

$$(3) \qquad \frac{\partial}{\partial x}(a_n u_{nx} + b_{1,n} u_{ny}) + \frac{\partial}{\partial y}(b_{2,n} u_{nx} + c_n u_{ny}) = 0$$

which coincides with H_n^* on R^* is analytic on $\overline{R}$† for each n; $\overline{M}$, $\overline{m}$, and G are independent of n. For each n, there exists a unique function $v_n(x, y)$, with $v_n(x_0, y_0) = 0$, which satisfies

$$(4) \qquad \begin{aligned} v_{nx} &= -(b_{2,n} u_{nx} + c_n u_{ny}), \\ v_{ny} &= a_n u_{nx} + b_{1,n} u_{ny}. \end{aligned}$$

Define $A_n(x, y)$, $B_n(x, y)$, $C_n(x, y)$ by

† This can be done by taking a sequence of regions $\overline{R}_n$, each bounded by analytic curves, which closes down on $\overline{R}$ and then assigning proper analytic boundary values on R_n^*. That the solutions of (3) exist under these conditions follows from the results stated in the article by L. Lichtenstein on the theory of elliptic partial differential equations in the Encyklopädie der Mathematischen Wissenschaften, vol. II 3^2, pp. 1280–1334.

$$
\begin{aligned}
A_n &= \frac{a_n^2 u_{nx}^2 + 2a_n b_{1,n} u_{nx} u_{ny} + (1 + b_{1,n}^2) u_{ny}^2}{a_n u_{nx}^2 + (b_{1,n} + b_{2,n}) u_{nx} u_{ny} + c_n u_{ny}^2}, \\
(5) \qquad B_n &= \frac{a_n b_{2,n} u_{nx}^2 + (a_n c_n + b_{1,n} b_{2,n} - 1) u_{nx} u_{ny} + b_{1,n} c_n u_{ny}^2}{a_n u^2{}_x + (b_{1,n} + b_{2,n}) u_{nx} u_{ny} + c_n u_{ny}}, \\
C_n &= \frac{(1 + b_{2,n}^2) u_{nx}^2 + 2b_{2,n} c_n u_{nx} u_{ny} + c_n^2 u_{ny}^2}{a_n u_{nx}^2 + (b_{1,n} + b_{2,n}) u_{nx} u_{ny} + c_n u_{ny}^2},
\end{aligned}
$$

where these expressions are defined; otherwise, let $A_n = C_n = 1$, $B_n = 0$. It is easily verified that A_n, B_n, C_n are measurable for each n and that

$$
\begin{aligned}
& A_n C_n - B_n^2 = 1, \ |A_n|, \ |B_n|, \ |C_n| \leqq K(\overline{M}, \overline{m}), \\
(6) \qquad & l \cdot (U_x^2 + U_y^2) \leqq A_n U_x^2 + 2B_n U_x U_y + C_n U_{ny}^2 \leqq L \cdot (U_x^2 + U_y^2), \\
& 0 < l \leqq L < \infty,
\end{aligned}
$$

where K, l, and L depend only on $\overline{M}$ and $\overline{m}$ (U is any function of class D_2 on R).

Let p, q, and r be three distinct points on R^* and π, κ, and ρ be three distinct points arranged in the same order on Σ^*, the boundary of the unit circle. Map $\overline{R}$ on $\overline{\Sigma}$ by functions $\xi = \xi_n(x, y)$, $\eta = \eta_n(x, y)$ where p, q, and r correspond respectively to π, κ, and ρ and where

$$
(7) \qquad \eta_{nx} = -(B_n \xi_{nx} + C_n \xi_{ny}), \qquad \eta_{ny} = A_n \xi_{nx} + B_n \xi_{ny}.\dagger
$$

Using the relations (4), (5), and (7) and Theorem 4, §2 we find that $u_n(\xi, \eta)$ and $v_n(\xi, \eta)$ are of class D_2 on $\overline{\Sigma}$ and satisfy $v_{n\xi} = -u_{n\eta}$, $v_{n\eta} = u_{n\xi}$ almost everywhere. They are therefore conjugate harmonic functions (in case $u_{nx} = u_{ny} = 0$, it is clear that $v_{nx} = v_{ny} = 0$ at almost all of these points; hence at almost all corresponding points, $v_{n\xi} = v_{n\eta} = u_{n\xi} = u_{n\eta} = 0$; this takes care of points for which $A_n = C_n = 1$, $B_n = 0$). Since the transformations (7) are equicontinuous both ways, a subsequence $\{n_k\}$ of the integers may be chosen so that the corresponding transformations and their inverses converge uniformly to a certain 1–1 continuous transformation of $\overline{R}$ into $\overline{\Sigma}$ and its inverse. Thus the sequences $\{u_{n_k}(\xi, \eta)\}$ and hence $\{u_{n_k}(x, y)\}$ converge uniformly to certain functions $u(\xi, \eta)$ and $u(x, y)$ respectively, $u(x, y)$ coinciding with H^* on R^* and $u(\xi, \eta)$ being harmonic; and the functions $v_{n_k}(\xi, \eta)$ converge uniformly on each closed subregion of Σ to $v(\xi, \eta)$, the conjugate of $u(\xi, \eta)$, and so

† It may be shown that $u = u_n(x, y)$, $v = v_n(x, y)$ carries $\overline{R}$ into a region on a finite sheeted Riemann surface. The transformation (7) merely amounts to mapping this Riemann region conformally on the unit circle $\overline{\Sigma}$ in a (ξ, η) plane. We shall use this transformation in the proof of Theorem 2 where this interpretation is not valid.

$v_{n_k}(x, y)$ converge uniformly on each closed subregion of R to a certain function $v(x, y)$.

It is easily verified that (see the proof of Theorem 4, §2)

$$\iint_D (A_n U_x^2 + 2B_n U_x U_y + C_n U_y^2) dx dy = \iint_{\Delta_n} (U_\xi^2 + U_\eta^2) d\xi d\eta \tag{8}$$

for any function U of class D_2 (in either (x, y) or (ξ, η)), D and Δ_n being corresponding regions of R and Σ. Since each u_n is harmonic, it follows that

$$\iint_\Sigma (u_{n\xi}^2 + u_{n\eta}^2) d\xi d\eta \leqq \iint_\Sigma (H_{n\xi}^{*2} + H_{n\eta}^{*2}) d\xi d\eta \leqq L \cdot G,$$

where L is the constant in (6) (using relations (6) and (8)). Hence, by (6) and (8), $D(u_n) \leqq L \cdot G/l$ (independent of n). Thus $u(x, y)$ and $v(x, y)$ are of class D_2 on $\overline{R}$ and R respectively. Furthermore, by Lemma 6, §1, it follows easily that

$$\iint_R [(v_x + b_2 u_x + c u_y)^2 + (v_y - a u_x - b_1 u_y)^2] dx dy = 0.$$

Thus the existence and properties (i), (ii), and (iii) are established.

To show (iv), define A, B, and C by (5) in terms of a, b_1, b_2, c, and u, and perform the transformation (7). The relations (6) hold with M and m replacing $\overline{M}$ and $\overline{m}$, and $u(x, y)$, $v(x, y)$ are carried into conjugate harmonic functions. By Theorem 4, §2, the functions $\xi(x, y)$ and $\eta(x, y)$ satisfy conditions $A[2\lambda, M(\alpha)]$ and $B[\lambda, N(\alpha)]$ where λ depends only on M and m and $M(\alpha)$ and $N(\alpha)$ depend only on $\overline{R}$, (p, q, r), (π, κ, ρ), M, and m. Thus it is easy to see how to complete the proof of (iv).

Now, suppose $\overline{U}$, $\overline{V}$ is another pair of functions obeying the conclusions (i), (ii), and (iii). Then U and V, where $U = \overline{U} - u$, $V = \overline{V} - v$, satisfy these conditions with $H^* \equiv 0$. By defining A, B, and C by (5) in terms of U, a, b_1, b_2, and c, and performing the transformation (7), we see that $U(\xi, \eta)$ and $V(\xi, \eta)$ are conjugate harmonic functions for which $U(\xi, \eta) = 0$ on Σ^* and $V(\xi_0, \eta_0) = 0$. Thus $U \equiv V \equiv 0$.

THEOREM 2. *Let $\overline{R}$ be a Jordan region and p, q, and r be distinct points on R^*; we assume that $(\overline{R}; p, q, r)$ satisfies a $D(\Lambda, d_0)$ condition. Let H^* be continuous on $\overline{R}$ and harmonic on R with $D(H^*)$ finite. Suppose a, b_1, b_2, c, d, and e satisfy* (2) *where d and e are not necessarily zero. Then the conclusions of Theorem* 1 *hold except that* (iv) *is replaced by*

(iv)$_a$ *u and v satisfy a condition $B[\lambda, N'(\alpha, \delta)]$ on R and $|u| \leqq P$ on R,*

(iv)$_b$ $$\iint_D (u_x^2 + u_y^2) dx dy \leqq N$$

for each closed subregion $\bar{D}$ of R; here λ depends only on M, m, Λ, and d_0, $N'(\alpha, \delta)$ and P depend only on these, on $(\bar{R}; p, q, r; \pi, \kappa, \rho)$, and on the maximum of $|H^*|$ on R^*, and N depends only on these and on the distance of $\bar{D}$ from R^*.

Proof. Approximate as in Theorem 1 to a, b_1, b_2, c, d, and e by sequences of analytic functions $\{a_n\}$, $\{b_{1,n}\}$, $\{b_{2,n}\}$, $\{c_n\}$, $\{d_n\}$, and $\{e_n\}$ and to H^* by harmonic functions H_n^* in such a way that if u_n is the solution of

$$\frac{\partial}{\partial x}(a_n u_{nx} + b_{1,n} u_{ny} + d_n) + \frac{\partial}{\partial y}(b_{2,n} u_{nx} + c_n u_{ny} + e_n) = 0$$

which coincides on R^* with H_n^*, then u_n is analytic on $\bar{R}$ and $D(H_n^*) \leqq G$, independent of n. Define A_n, B_n, and C_n by (5) and perform the transformation (7). We find as in Theorem 1 that u_n and v_n are of class D_2 in (ξ, η) on $\bar{\Sigma}$ and satisfy

$$v_{n\xi} = -(u_{n\eta} + E_n), \qquad v_{n\eta} = u_{n\xi} + D_n \tag{9}$$

almost everywhere, where

$$D_n = d_n y_{n\eta} - e_n x_{n\eta}, \qquad E_n = -(d_n y_{n\xi} - e_n x_{n\xi}). \tag{10}$$

Also

$$A_n C_n - B_n^2 = 1, \ |A_n|, \ |B_n|, \ |C_n| \leqq K(\bar{M}, \bar{m}),$$

$$l\cdot(U_x^2 + U_y^2) \leqq A_n U_x^2 + 2B_n U_x U_y + C_n U_y^2 \leqq L\cdot(U_x^2 + U_y^2),$$
$$0 < l \leqq L < \infty,$$

$$\iint_{C(P,\rho)\cdot\Sigma} (D_n^2 + E_n^2)\,d\xi d\eta \leqq N_1(\bar{M}, \bar{m}, \Lambda, d_0)\cdot\rho^{2\lambda}, \lambda > 0, \lambda = \lambda(\bar{M}, \bar{m}, \Lambda), \tag{11}$$

$$\iint_D (A_n U_x^2 + 2B_n U_x U_y + C_n U_y^2)\,dxdy = \iint_{\Delta_n} (U_\xi^2 + U_\eta^2)\,d\xi d\eta,$$

where U stands for any function of class D_2, and D and Δ_n denote corresponding regions of R and Σ respectively, and K, N_1 and λ depend only on the quantities indicated, and l and L depend only on $\bar{M}$ and $\bar{m}$. These results are obtained by straightforward computation, the use of Theorems 3 and 4 of §2, and the relations of Theorem 1.

Thus

$$u_n = \bar{H}_n(\xi, \eta) + u_{0,n}(\xi, \eta), \tag{12}$$

where $u_{0,n}$ has the significance of §4 and $\bar{H}_n(\xi, \eta)$ is the function harmonic on $\bar{\Sigma}$, and taking on the boundary values of $H_n^*(\xi, \eta)$. Now, as in Theorem 1, we see that $D(u_n) \leqq L\cdot G/l$ and that a subsequence $\{u_{n_k}\}$ converges uniformly on $\bar{R}$ to a function $u(x, y)$ and $\{v_{n_k}\}$ converges to a function $v(x, y)$ uniformly

on each closed subregion of R. Thus u and v are of class D_2 and, as in Theorem 1, are seen to satisfy (1) almost everywhere. It is clear that (i), (ii), and (iii) are fulfilled. That the pair (u, v) is unique follows from Theorem 1, since if u' and v' were another pair with the same boundary values, $u'-u$ and $v'-v$ would satisfy (1) with $d=e=0$ and obey the conclusions (i), (ii), and (iii) of Theorem 1 with $u'-u=0$ on R^* and $v'(x_0, y_0)-v(x_0, y_0)=0$. That $|u_n| \leqq \overline{P}$ which depends only on $\overline{M}$, $\overline{m}$, Λ, d_0, and the maximum of $|H^*|$ on R^* follows from Theorems 1 and 2 of §3.

Using Theorem 3, §2, and (12), we see that (iv) holds except that $\overline{M}$ and $\overline{m}$ intervene instead of M and m. Using (11), we see that

$$\begin{aligned}\iint_D (u_x^2 + u_y^2)dxdy &\leqq \liminf_{n\to\infty} \iint_D (u_{nx}^2 + u_{ny}^2)dxdy \\ &\leqq \frac{1}{l}\liminf_{n\to\infty} \iint_D (A_n u_{nx}^2 + 2B_n u_{nx}u_{ny} + C_n u_{ny}^2)dxdy \\ &= \frac{1}{l}\liminf_{n\to\infty} \iint_{\Delta_n} (u_{n\xi}^2 + u_{n\eta}^2)d\xi d\eta \leqq N,\end{aligned}$$

where N depends on the quantities indicated in the theorem except that $\overline{M}$ and $\overline{m}$ intervene instead of M and m; this is true since

$$\iint_\Sigma (u_{n,0\xi}^2 + u_{n,0\eta}^2)d\xi d\eta \leqq \iint_\Sigma (D_n^2 + E_n^2)d\xi d\eta$$

and regions $\bar{D}$ at a distance $\geqq \delta$ from R^* are carried into regions $\overline{\Delta}_n$ at a distance $\geqq m(\delta)$ from Σ^*, where $m(\delta)$ depends only on $\overline{M}$, $\overline{m}$, and $(\overline{R}; p, q, r; \pi, \kappa, \rho)$ and in such a region $|\overline{H}_{n\xi}|$ and $|\overline{H}_{n\eta}| \leqq 2h_n\pi^{-1}[m(\delta)]^{-1}$ where h_n denotes the maximum of $|H_n^*|$ on R^*. To get rid of $\overline{M}$ and $\overline{m}$, merely define A, B, C in terms of a, b_1, b_2, c, and u and perform (7); all the conclusions then hold with M and m replacing $\overline{M}$ and $\overline{m}$.

THEOREM 3. *Let $\{a_n\}$, $\{b_{1,n}\}$, $\{b_{2,n}\}$, $\{c_n\}$, $\{d_n\}$, and $\{e_n\}$ be sequences of measurable functions which converge to a, b_1, b_2, c, d, and e respectively almost everywhere and which satisfy*

$$|a_n|,\ |b_{1,n}|,\ |b_{2,n}|,\ |c_n|,\ |d_n|,\ |e_n| \leqq M,$$
$$4a_nc_n - (b_{1,n} + b_{2,n})^2 \geqq m > 0, \qquad a_n > 0.$$

For each n, let u_n and v_n be of class D_2 on $\overline{R}$ and R respectively and satisfy

$$v_{nx} = -(b_{2,n}u_{nx} + c_nu_{ny} + e_n), \qquad v_{ny} = a_nu_{nx} + b_{1,n}u_{ny} + d_n$$

almost everywhere with $|u_n| \leqq G$ on R^*(M, m, and G being independent of n). Then

(i) $|a|, |b_1|, |b_2|, |c|, |d|, |e| \leqq M$, $4ac-(b_1+b_2)^2 \geqq m>0$, $a>0$,

(ii) *$\{u_n\}$ and $\{v_n\}$ are uniformly bounded and satisfy uniform Hölder conditions on each closed subregion $\bar{D}$ of R, which depend only on M, m, $\bar{R}$, G, and the distance $\delta(\bar{D})$ of $\bar{D}$ from R^*,*

(iii) $\iint_D (u_{nx}^2+u_{ny}^2+v_{nx}^2+v_{ny}^2)dx\,dy \leqq P[M, m, G, \bar{R}, \delta(\bar{D})]$, *and*

(iv) *if the subsequences $\{u_{n_k}\}$, $\{v_{n_k}\}$ are chosen to converge uniformly on each closed subregion $\bar{D}$ of R to functions u and v, then u and v satisfy* (ii) *and* (iii) *and*

$$v_x = -(b_2u_x + cu_y + e), \qquad v_y = au_x + b_1u_y + d,$$

almost everywhere on R. The same conclusions hold if we merely assume that each u_n is of class D_2 on R with $|u_n| \leqq G$.

Proof. (i) is obvious and (ii) and (iii) have been proved in Theorem 2. To prove (iv), let $\bar{D}$ be a closed subregion of R. The conclusions (i), (ii), and (iii) hold for $\bar{D}$ and $\{u_{n_k}\}$ and $\{v_{n_k}\}$ converge uniformly on $\bar{D}$ to u and v which are of class D_2 on $\bar{D}$. Then, by Lemma 6, §1,

$$0 \leqq \iint_D [(v_x + b_2u_x + cu_y + e)^2 + (v_y - au_x - b_1u_y - d)^2 dxdy$$

$$\leqq \liminf_{n\to\infty} \iint_D [(v_{nx} + b_{2,n}u_{nx} + c_nu_{ny} + e_n)^2 + (v_{ny} - a_nu_{nx} - b_{1,n}u_{ny} - d_n)^2]dxdy = 0.$$

THEOREM 4. *If a, b_1, b_2, c, d, and e satisfy hypotheses* (2) *and satisfy*

$$|\phi(x_1, y_1) - \phi(x_2, y_2)| \leqq Nr^\lambda, \quad \lambda > 0, \quad r = [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2},$$

($\phi(x, y)$ standing for a, b_1, b_2, c, d, or e) on $\bar{R}$, then if $u(x, y)$ and $v(x, y)$ constitute a solution of (1) *with $|u| \leqq G$, then u_x, u_y, v_x, and v_y satisfy a uniform Hölder condition with the same exponent λ on any subregion $\bar{D}$ of R where the constant N' depends on M, m, N, λ, $\bar{R}$, and the distance of $\bar{D}$ from R.*†

5. **Applications to the calculus of variations.** In this section, we shall discuss the differentiability properties of a function $z(x, y)$ which minimizes

$$\iint_R f(x, y, z, p, q)dxdy \qquad (p = z_x, q = z_y) \tag{1}$$

among all functions having the same boundary values. We shall assume that f is continuous together with its first and second partial derivatives for all values of (x, y, z, p, q) and that the second derivatives satisfy a uniform Hölder condition on each bounded portion of (x, y, z, p, q) space, with

† See E. Hopf, loc. cit.

$$f_{pp}f_{qq} - f_{pq}^2 > 0, \qquad f_{pp} > 0,$$

everywhere.

LEMMA 1.† *If $z(x, y)$ is a solution of* (1) *on a region $\bar{R}$, which solution satisfies a uniform Lipschitz condition, then*

$$\int_{D^*} f_p dy - f_q dx = \iint_D f_z dx dy,$$

for almost all rectangles D in R.

Proof. Let $\zeta(x, y)$ be any function which satisfies a Lipschitz condition on $\bar{R}$ and which vanishes on R^*. Then it follows in the usual way that

$$\iint_R (\zeta f_z + \zeta_x f_p + \zeta_y f_q) dx dy = 0.$$

Now let $\bar{D}: a \leqq x \leqq b,\ c \leqq y \leqq d$, be any closed rectangle in R, and define

$$\phi(x, y) = \int_a^x f_z[\xi, y, z(\xi, y), p(\xi, y), q(\xi, y)] d\xi.$$

Then, if ζ is any function satisfying a uniform Lipschitz condition on $\bar{R}$ and zero outside and on D^*, we see that

$$\iint_D (\zeta\phi_x + \zeta_x f_p + \zeta_y f_q) dx dy = \iint_D \zeta_x(f_p - \phi) + \zeta_y f_q) dx dy = 0.$$

Thus, it follows from a theorem of A. Haar‡ that

$$\int_{\Delta^*} (f_p - \phi) dy - f_q dx = 0$$

for almost all rectangles $\bar{\Delta}$ of D, and hence that

$$\int_{\Delta^*} f_p dy - f_q dx = \int_{\Delta^*} \phi dy = \iint_\Delta f_z dx dy.$$

Since R may be written as the sum of a denumerable number of rectangles D, the lemma follows.

---

† This lemma is equivalent to Haar's equations for a minimizing function in a double integral problem, first stated and proved by him in the case that p and q are continuous. See A. Haar, *Über die Variation der Doppelintegrale*, Journal für die Reine und Angewandte Mathematik, vol. 149 (1919), pp. 1–18.

‡ A. Haar, *Über das Plateausche Problem*, Mathematische Annalen, vol. 97 (1926–27), pp. 124–158, particularly pp. 146–151.

THEOREM 1. *If $z(x, y)$ satisfies the hypotheses of Lemma* 1, *then it is continuous together with its first and second partial derivatives and the latter satisfy uniform Hölder conditions on each closed subregion of R. If $f(x, y, z, p, q)$ is analytic, $z(x, y)$ is also.*

Proof. Let h be a small rational number and let (x_0, y_0) be an interior point of R. We define the region R_h as the set of all points (x_1, y_1) of R such that (1) the segment $x_1 \leqq x \leqq x_1+h$, $y=y_1$ if $h>0$ or $x_1+h \leqq x \leqq x_1$, $y=y_1$ if $h<0$ lies in R and (2) (x_1, y_1) may be joined to (x_0, y_0) by a curve, all of whose points satisfy (1). If $|h| \leqq h_0$, R_h is a non-vacuous Jordan region (h_0 rational). Let $\bar{D}$ be a closed subregion of R; for $|h| < h_1$, R_h contains $\bar{D}$. Also, if H is any rectangle out of a certain set of "almost all" rectangles of D, we have

$$\begin{aligned}
&\int_{H^*} \{f_p[x + h, y, z(x + h, y), p(x + h, y), q(x+h, y)]dy \\
(2) \qquad &\qquad - f_q[x + h, y, z(x + h, y), p(x + h, y), q(x + h, y)]dx\} \\
&= \iint_H f_z[x + h, y, z(x + h, y), p(x + h, y), q(x + h, y)]dxdy,
\end{aligned}$$

$$\begin{aligned}
&\int_{H^*} \{f_p[x, y, z(x, y), p(x, y), q(x, y)]dy \\
(3) \qquad &\qquad - f_q[x, y, z(x, y), p(x, y), q(x, y)]dx\} \\
&= \iint_H f_z[x, y, z(x, y), p(x, y), q(x, y)]dxdy,
\end{aligned}$$

for all rational h with $|h| < h_1$.

Let

$$z_h = \frac{1}{h}\int_x^{x+h} z(\xi, y)d\xi.$$

Then

$$p_h(x, y) = \frac{z(x + h, y) - z(x, y)}{h}, \quad q_h(x, y) = \frac{1}{h}\int_x^{x+h} q(\xi, y)d\xi,$$

$$p_{hx} = \frac{p(x + h, y) - p(x, y)}{h}, \qquad p_{hy} = \frac{q(x + h, y) - q(x, y)}{h} = q_{hx},$$

almost everywhere ($|h| < h_1$). Clearly p_h satisfies a Lipschitz condition on $\bar{D}$ for each h with $0 < |h| < h_1$, and hence is of class D_2. Also $|p_h|$ is bounded independently of h.

Subtracting (3) from (2) and dividing by h, we find that

$$\int_{H^*} \{[a_h(x, y)p_{hx} + b_h(x, y)p_{hy} + d_h(x, y)]dy \tag{4}$$
$$- [b_h(x, y)p_{hx} + c_h(x, y)p_{hy} + e_h(x, y)]dx\} = 0$$

on almost all rectangles H of D (for each rational h with $|h| < h_1$). Here we may take

$$a_h(x, y) = \int_0^1 f_{pp}\{x + th, y, z(x, y) + t[z(x + h, y) - z(x, y)], p(x, y)$$
$$+ t[p(x + h, y) - p(x, y)], q(x, y) + t[q(x + h, y) - q(x, y)]\}dt,$$

$$b_h(x, y) = \int_0^1 f_{pq}dt, \qquad c_h(x, y) = \int_0^1 f_{qq}dt,$$

$$e_h(x, y) = \int_0^1 f_{qx}dt + \frac{z(x + h, y) - z(x, y)}{h}\int_0^1 f_{qz}dt,$$

$$d_h(x, y) = \int_0^1 f_{px}dt + \frac{z(x + h, y) - z(x, y)}{h}\int_0^1 f_{pz}dt$$
$$- \frac{1}{h}\int_x^{x+h} f_z[\xi, y, z(\xi, y), p(\xi, y), q(\xi, y)]d\xi,$$

where the arguments which are not indicated in b_h, c_h, d_h, and e_h are all the same as that in a_h. Clearly $|a_h|$, $|b_h|$, $|c_h|$, $|d_h|$, and $|e_h|$ are uniformly bounded for $|h| < h_1$ and it can easily be shown as in the proof of Lemma 2, §2, that

$$a_h c_h - b_h^2 \geqq m > 0$$

for all h with $|h| < h_1$. Also from (4), it follows that, for each such h, there exists a function $v_h(x, y)$ which satisfies a Lipschitz condition on $\overline{R}$ and which satisfies

$$v_h(x_0, y_0) = 0, \qquad v_{hx} = -(b_h u_{hx} + c_h u_{hy} + e_h), \qquad v_{hy} = a_h u_{hx} + b_h u_{hy} + d_h$$

almost everywhere. Also, if $\{h_n\} \to 0$ (all rational with $|h_n| < h_1$) $\{a_{h_n}\}$, $\{b_{h_n}\}$, $\{c_{h_n}\}$, $\{d_{h_n}\}$, and $\{e_{h_n}\}$ tend to $f_{pp}[x, y, z(x, y), p(x, y), q(x, y)]$, f_{pq}, f_{qq}, $f_{px} + f_{pz}\cdot p - f_z$, and $f_{qx} + f_{qz}\cdot p$ respectively. Hence, from Theorem 3, §4 and the fact that p_{h_n} tends to p almost everywhere, it follows that $p(x, y)$ satisfies a uniform Hölder condition on each closed subregion $\overline{\Delta}$ of D. Similarly it may be shown that $q(x, y)$ also satisfies the same type of condition.

If we choose a subsequence so $\{v_{h_n}\} \to v$, then we have that

$$v_x = -(f_{pq}p_x + f_{qq}q_x + [f_{qx} + f_{qz}p]), \; v_y = f_{pp}p_x + f_{pq}p_y + [f_{px} + f_{pz}p - f_z]$$

almost everywhere. From Theorem 4, §4 it follows that p_x and p_y satisfy

uniform Hölder conditions on each subregion of R; the same statement holds for q_x and q_y. This proves the theorem. The last statement has been shown to hold by E. Hopf† if it is known that $p(x, y)$ and $q(x, y)$ satisfy Hölder conditions which fact we proved above.

6. **Applications to quasi-linear elliptic equations.** In this section, we shall consider the equations of the form

$$a(x, y)z_{xx} + 2b(x, y)z_{xy} + c(x, y)z_{yy} = d(x, y), \tag{1}$$

$$a(x, y, z, p, q)z_{xx} + 2b(x, y, z, p, q)z_{xy} + c(x, y, z, p, q)z_{yy} = \lambda d(x, y, z, p, q), \tag{2}$$

in which we assume that the functions a, b, c, and d are defined and continuous for all values of their arguments with $ac-b^2=1$, $a>0$, and that these functions satisfy a uniform Hölder condition in each bounded portion of the space in which they are defined. Then it is known‡ that there exists a solution of (1) which is defined in the unit circle $\overline{\Sigma}$ and vanishes on Σ^* and that its second derivatives satisfy a uniform Hölder condition on each closed subregion of Σ. A more precise statement is given in Lemma 1 below.

Lemma 1. *Let $z(x, y)$ be the solution of* (1) *which vanishes on Σ^*, let k be the maximum of $|d|/(a+c)$ on $\overline{\Sigma}$, and let l be the maximum of $(a+c)$ on $\overline{\Sigma}$. Then*

$$|z| \leqq \frac{k}{2}, \quad |p|, \quad |q| \leqq 120l^3k$$

on Σ, and p and q satisfy uniform Hölder conditions on each closed subregion $\overline{\Delta}$ of Σ which depends only on k and l and the distance of $\overline{\Delta}$ from Σ^*, and z_{xx}, z_{xy}, and z_{yy} satisfy uniform Hölder conditions on each closed subregion $\overline{\Delta}$ of Σ which depend only on the above and on the Hölder conditions satisfied by a, b, c, and d.

Proof. First, suppose that a, b, c, and d are analytic on $\overline{\Sigma}$ so that z is analytic on $\overline{\Sigma}$. It is known§ that if $d_1(x, y) \leqq d(x, y) \leqq d_2(x, y)$ on $\overline{R}$ and if z_1 and z_2 are the corresponding solutions of (1), then $z_1 \geqq z \geqq z_2$. Hence if we choose

$$z_1 = k(1 - x^2 - y^2)/2, \qquad z_2 = -z_1,$$
$$d_1(x, y) = -k(a + c), \qquad d_2(x, y) = k(a + c),$$

we see that $k/2 \geqq z \geqq -k/2$ on $\overline{\Sigma}$. Also, since $z_1 - z \geqq 0$ and $z - z_2 \geqq 0$ on $\overline{\Sigma}$, we see that

† Loc. cit.

‡ For it is known (see Lichtenstein, loc. cit.) that the solution exists if a, b, c, and d are analytic and the result follows from Theorem 4, §4 by approximations.

§ See for instance in S. Bernstein, loc. cit., second paper.

$$\frac{\partial z_2}{\partial r} = k \geqq \frac{\partial z}{\partial r} \geqq -k = \frac{\partial z_1}{\partial r}$$

on Σ^*. Since $\partial z/\partial\theta = 0$ on Σ^*, it follows that

$$|p|, |q| \leqq k \text{ on } \Sigma^*.$$

Now p and q satisfy the equations

$$ap_x + 2bp_y + cq_y = d, \qquad p_y - q_x = 0.$$

If we set $u = -p$, $v = q$, and $v = p$, $u = q$, the equations become

$$v_x = -u_y, \qquad v_y = \frac{a}{c}u_x + \frac{2b}{c}u_y + \frac{d}{c}, \tag{3}$$

$$v_x = -\left(\frac{2b}{a}u_x + \frac{c}{a}u_y - \frac{d}{a}\right), \qquad v_y = u_x, \tag{4}$$

respectively, each of which is a system of the type treated in §4. Thus, by Theorem 3, §4, we see that p and q are bounded and satisfy uniform Hölder conditions as desired in the lemma, and from Theorem 4, §4, we see that z_{xx}, z_{xy}, and z_{yy} satisfy the desired conditions; these conditions hold whether a, b, c, and d are analytic or not.

To see what the bounds are for p and q, let $\overline{R}$ be the unit circle and let π, κ, ρ and p, q, r be three equally spaced points on Σ^* and R^* respectively. Clearly $(\overline{\Sigma};\ \pi,\ \kappa,\ \rho)$ and $(\overline{R};\ p,\ q,\ r)$ satisfy $D(L, d_0)$ conditions where $L = 4\pi\cdot 3^{-3/2}$ and $d_0 = 3^{1/2}$. In making the transformation (7) of §4, we see that the ratio of the maximum to the minimum magnification at each point is given by $\mu + (\mu^2 - 1)^{1/2}$ where

$$\mu = \frac{A + C}{2} = \frac{1}{2}\cdot\frac{\left(1 + \frac{a^2}{c^2}\right)u_x^2 + \frac{4ab}{c^2}u_xu_y + \left(2 + \frac{4b^2}{c^2}\right)u_y^2}{\frac{a}{c}u_x^2 + \frac{2b}{c}u_xu_y + u_y^2}$$
$$\leqq \frac{(a + c)^2}{2} - 1$$

since $a + c \geqq 2$. Thus the K of Theorem 2, §2 is $l^2 - 2$, and hence

$$\iint_{C(P_0,r)\cdot R} (x_\xi^2 + x_\eta^2 + y_\xi^2 + y_\eta^2)d\xi d\eta \leqq 2\pi(l^2 - 2)(3^{-1/2}\cdot 2r)^{2/[(l^2-2)(L+1)]}.$$

Since, on R, $D = (d/c)y_\eta$, $E = -(d/c)y_\xi$ in (3) and $D = (d/a)x_\eta$, $E = -(d/a)x_\xi$ in (4) (using relations (10) of §4), it follows that

$$\iint_{C(P,\rho)\cdot R} (D^2 + E^2)d\xi d\eta \leqq \max_{(x,\,y)\text{ on }\bar{\Sigma}} \frac{4\pi(l^2-2)}{3^{1/2}}\cdot\frac{d^2}{c^2}\,\rho^{2\lambda},$$

$$\iint_{C(P,\rho)\cdot R} (D^2 + E^2)d\xi d\eta \leqq \max_{(x,\,y)\epsilon\bar{\Sigma}} \frac{4\pi(l^2-2)}{3^{1/2}}\cdot\frac{d^2}{a^2}\cdot\rho^{2\lambda},\quad \lambda = \frac{1}{(l^2-2)(L+1)^2},$$

in cases (3) and (4) respectively. Let us call the first $\gamma\rho^{2\lambda}$ and the second $\alpha\rho^{2\lambda}$. Referring to Theorem 1, §3 and using the notations of Theorems 2 and 3 and the remark, we see that

$$|U_1|,\ |V_1| \leqq 4(1+[2\lambda]^{-1})\left(\frac{\gamma}{2\pi}\right)^{1/2},\quad |U_1|,\ |V_1| \leqq 4(1+[2\lambda]^{-1})\left(\frac{\alpha}{2\pi}\right)^{1/2}$$

in cases (3) and (4) respectively so that

$$|u_{0,1}|,\ |v_{0,1}| \leqq 8(1+[2\lambda]^{-1})\left(\frac{\gamma}{2\pi}\right)^{1/2},\ |u_{0,1}|,\ |v_{0,1}| \leqq 8(1+[2\lambda]^{-1})\left(\frac{\alpha}{2\pi}\right)^{1/2}$$

in (3) and (4) respectively, so that, finally

$$|u|,\ |v| \leqq 8\left(1+\frac{1}{2\lambda}\right)\left(\frac{\gamma}{2\pi}\right)^{1/2} + k,\ |u|,\ |v| \leqq 8\left(1+\frac{1}{2\lambda}\right)\left(\frac{\alpha}{2\pi}\right)^{1/2} + k,$$

in (3) and (4) respectively. Since $|u|$, $|v|$ are merely $|p|$, $|q|$ in one order or the other, and since $2|d|/(a+c)$ is between $|d|/c$ and $|d|/a$, we may substitute $4k^2$ for d^2/a^2 and d^2/c^2 so that we obtain

$$|p|,\ |q| \leqq 8[1+\tfrac{1}{2}(l^2-2)\cdot(4\pi\cdot 3^{-3/2}+1)^2]2^{-1/2}\cdot 3^{-1/4}\cdot l\cdot 2k \leqq 120l^3k$$

since $l \geqq 1$.

Definition 1. A function $z^*(x, y)$ is said to satisfy a *three point condition with constant* Δ on Σ^* if for each plane $z = ax+by+c$ which passes through three points of the curve $z = z^*(x, y)$, (x, y) on Σ^*, we have $a^2+b^2 \leqq \Delta^2$.

Lemma 2. *Let $z^*(x, y)$ satisfy a three point condition with constant Δ on Σ^*, and let z be the solution of* (1) *which takes on these boundary values, $d(x, y)$ being assumed to be identically zero on $\bar{\Sigma}$. Then*

(i) $\min_{(x,\,y)\text{ on }\Sigma^*} z^* \leqq z(x, y) \leqq \max_{(x,\,y)\epsilon\Sigma^*} z^*$, $(x, y)\epsilon\bar{\Sigma}$ *in* $z(x, y)$, *and*

(ii) $p^2 + q^2 \leqq \Delta^2$.

Proof. This is well known.†

Theorem 1. *Let $z^*(x, y)$ satisfy a three point condition with constant Δ on*

† See, for instance, J. Schauder, *Über das Dirichletsche Problem im Grossen für nichtlineare elliptische Differentialgleichungen*, Mathematische Zeitschrift, vol. 37 (1933), pp. 623–634.

Σ^* and let us assume that $\lambda=0$ in equation (2). *Then there exists a solution $z(x, y)$ of* (2) *which coincides with z^* on Σ^*.*

Proof. Let M denote the maximum of $|z^*|$ on Σ^* and let l be the least upper bound of $a(x, y, z, p, q)+c(x, y, z, p, q)$ for all functions $z(x, y)$ which satisfy Lipschitz conditions and for which $|z| \leqq M$, $p^2+q^2 \leqq \Delta^2$.

Now, let $z_0(x, y)$ be the harmonic function taking on the given boundary values and define functions $z_n(x, y)$ for $n \geqq 1$ as the solutions of

$$a(x, y, z_n, p_n, q_n)z_{n+1,xx} + 2b(x, y, z_n, p_n, q_n)z_{n+1,xy} + c(x, y, z_n, p_n, q_n)z_{n+1,yy} = 0$$

which coincide with z^* on Σ^*. At each stage

$$|z_n| \leqq M, \qquad p_n^2 + q_n^2 \leqq \Delta^2$$

and a_n, b_n, c_n satisfy Hölder conditions which depend only on Δ, M, and l, where a_n stands for $a(x, y, z_n, p_n, q_n)$, for instance. Since p_n and q_n satisfy the equations

$$a_{n-1}p_{nx} + 2b_{n-1}p_{ny} + c_{n-1}q_{ny} = 0, \qquad p_{ny} - q_{nx} = 0$$

and $|p_n|$, $|q_n| \leqq \Delta$, it follows from §4 that p_n and q_n satisfy uniform Hölder conditions on each closed subregion $\bar{D}$ of Σ which depend only on l, Δ, and the distance of $\bar{D}$ from Σ^* (and not on n). Thus, by Theorem 4, § 4, $z_{n,xx}$, $z_{n,xy}$, and $z_{n,yy}$ satisfy similar conditions, since a_{n-1}, b_{n-1}, and c_{n-1} satisfy uniform Hölder conditions on subregions of Σ which depend only on the above quantities. Thus a subsequence $\{n_k\}$ may be chosen so that the sequence $\{z_{n_k}\}$ converges uniformly on $\bar{\Sigma}$ to a function z and so that the sequences $\{z_{n_k,x}\}$, $\{z_{n_k,y}\}$, $\{z_{n_k,xx}\}$, $\{z_{n_k,xy}\}$, and $\{z_{n_k,yy}\}$ converge uniformly on each closed subregion of Σ to the corresponding derivatives of z. Clearly z is the desired solution.

THEOREM 2. *If $|\lambda|$ is sufficiently small, the equation* (2) *possesses a solution which vanishes on Σ^*.*

Proof. Let $l(z)$ and $k(z)$ denote the least upper bounds of

$$a[x, y, z(x, y), p(x, y), q(x, y)] + c[x, y, z(x, y), p(x, y), q(x, y)],$$
$$\frac{d[x, y, z(x, y), p(x, y), q(x, y)]}{a[x, y, z(x, y), p(x, y), q(x, y)] + c[x, y, z(x, y), p(x, y), q(x, y)]}$$

respectively. For all z with $z^2+p^2+q^2 \leqq \alpha^2$, $l(z) \leqq l$, $k(z) \leqq k$. Now if $z(x, y)$ is a function for which $z^2+p^2+q^2 \leqq \alpha^2$ and for which $p(x, y)$, $q(x, y)$ satisfy a uniform Hölder condition on each closed subregion of Σ, the solution Z of

$$a(x, y, z, p, q)Z_{xx} + 2b(x, y, z, p, q)Z_{xy} + c(x, y, z, p, q)Z_{yy} = \lambda d(x, y, z, p, q)$$

which vanishes on Σ^* is such that Z_{xx}, Z_{xy}, Z_{yy} satisfy Hölder conditions on each closed subregion $\bar{D}$ of Σ which depend only on the Hölder conditions satisfied by $a, b, c, \lambda d$ and z, p, and q, and Z_x and Z_y satisfy Hölder conditions which depend only on α, k, l, and λ. Also, by Lemma 1

$$(Z^2 + P^2 + Q^2)^{1/2} \leqq 120 \cdot 3^{1/2} \cdot l^3 \cdot k \cdot |\lambda|$$

which is $\leqq \alpha$ if λ is small enough. The successive approximations may be carried through as in Theorem 1.

THEOREM 3. *Let $L(G)$ and $K(G)$ denote the least upper bounds of $l(z)$ and $k(z)$, respectively, for all z with $z^2+p^2+q^2 \leqq G^2$. If, in addition to our hypotheses, we assume that*

$$\lim_{G\to\infty} \frac{L^3(G)\cdot K(G)}{G} = 0,$$

the equation (2) *possesses a solution on Σ which vanishes on Σ^* for each value of λ.*

Proof. For each $\mu > 0$, there exists a number N_μ such that

$$120 \cdot 3^{1/2} \cdot L^3(G) \cdot K(G) \leqq \mu G$$

for all $G \geqq N_\mu$. Let M_μ be the least upper bound of $120 L^3(G) K(G) \cdot 3^{1/2}$ for all $G \leqq N_\mu$. If we let P_μ be the larger of N_μ and $\mu^{-1} \cdot M_\mu$, we see that

$$120 l^3(z) \cdot k(z) \cdot 3^{1/2} \leqq \mu P_\mu$$

for all z for which $z^2+p^2+q^2 \leqq P_\mu^2$. Thus, if z is such a function and Z is the solution of

$$a(x, y, z, p, q)Z_{xx} + 2b(x, y, z, p, q)Z_{xy} + c(x, y, z, p, q)Z_{yy} = \lambda d(x, y, z, p, q)$$

which vanishes on Σ^*, then

$$(Z^2 + P^2 + Q^2)^{1/2} \leqq |\lambda| \cdot \mu \cdot P_\mu$$

which is $\leqq P_\mu$ if $|\lambda| \leqq \mu^{-1}$. The remainder of the proof proceeds as in Theorems 1 and 2.

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIF.

A CLASS OF POLYNOMIALS*

BY

LEONARD CARLITZ

1. **Introduction.** For an indeterminate x in the $GF(p^n)$, put

$$[k] = x^{p^{nk}} - x, \quad F_k = [k][k-1]^{p^n} \cdots [1]^{p^{n(k-1)}}, \quad F_0 = 1; \tag{1.1}$$

then we define the function† $\psi(t)$ by means of

$$\psi(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{F_k} t^{p^{nk}}, \tag{1.2}$$

where t takes on the values

$$t = \sum_{j=0}^{\infty} c_{m-j} x^{m-j} \qquad (c_j \text{ in } GF(p^n)).$$

Then $\psi(t)$ has the linearity properties

$$\psi(t+u) = \psi(t) + \psi(u), \qquad \psi(ct) = c\psi(t), \tag{1.3}$$

for arbitrary c in $GF(p^n)$; further from (1.2) it follows that

$$-\psi(xt) = \psi^{p^n}(t) - x\psi(t). \tag{1.4}$$

In turn (1.4) implies the general relation

$$(-1)^m \psi(Mt) = \omega_M(\psi(t)), \tag{1.5}$$

where M is a polynomial in $GF(p^n)$ of degree m in x, and

$$\omega_M(u) = \sum_{j=0}^{m} \frac{(-1)^{m-j}}{F_j} \psi_j(M) u^{p^{nj}}. \tag{1.6}$$

It remains to define $\psi_j(t)$. We put

$$\begin{bmatrix} k \\ i \end{bmatrix} = \frac{F_k}{F_i L_{k-i}^{p^{ni}}}, \qquad \begin{bmatrix} k \\ 0 \end{bmatrix} = \frac{F_k}{L_k}, \qquad \begin{bmatrix} k \\ k \end{bmatrix} = 1,$$

for F_k as defined in (1.1), and

$$L_k = [k][k-1] \cdots [1], \qquad L_0 = 1;$$

then we have

* Presented to the Society, December 29, 1936; received by the editors November 23, 1936.

† For a discussion of $\psi(t)$ and $\psi_k(t)$ see the Duke Mathematical Journal, vol. 1 (1935), pp. 137–168.

$$(1.7)\qquad \psi_k(t) = \sum_{i=0}^{k} (-1)^{k-i} \begin{bmatrix} k \\ i \end{bmatrix} t^{p^{ni}}, \qquad \psi_0(t) = t.$$

In this paper we shall be interested first in the polynomials $\omega_M(u)$. Evidently (1.5) implies

$$(1.8)\qquad \omega_{MN}(u) = \omega_M(\omega_N(u)),$$

for arbitrary polynomials M, N. Assume next that M is *primary*, that is, the coefficient of the highest power of x occurring in M is the unit element of $GF(p^n)$. Then we define a class of polynomials $W_M(u)$ related to $\omega_M(u)$ by means of

$$(1.9)\qquad \omega_M(u) = \prod_{A|M} W_A(u),$$

the product extending over all (primary) polynomials A dividing M.

As we shall see below, the polynomials $W_M(u)$ have many properties analogous to those of the well-known cyclotomic polynomials.* In particular $W_M(u)$ is irreducible in the ring $F[u]$, where $F = F(x, p^n)$ is the field of rational functions of x with coefficients in $GF(p^n)$. Again if P is an irreducible polynomial in x, the factorization of $W_M(u)$ (mod P) is determined by a very simple rule. For example, if $P \nmid M$, define $e > 0$ as the smallest exponent such that

$$P^e \equiv 1 \pmod{M},$$

and put $\phi(M) = er$, where $\phi(M)$ is the Euler function for polynomials M; then we have the factorization

$$W_M(u) \equiv f_1(u) f_2(u) \cdots f_r(u) \pmod{P},$$

where each $f_i(u)$ is irreducible (mod P) and of degree e in u. Applications are made to the congruence $\omega_M(u) \equiv \delta$ (mod P).

2. **Notation; properties of $\omega_M(u)$.** It will be convenient to fix certain notation. If $GF(p^n)$ denotes a fixed Galois field of order p^n, we denote by $R = R(x, p^n)$ the ring of polynomials in the indeterminate x with coefficients in $GF(p^n)$. Similarly $F = F(x, p^n)$ denotes the field of rational functions of x with coefficients in $GF(p^n)$. For an additional indeterminate u, $R[u]$ and $F[u]$ denote rings of polynomials in u with coefficients in R and F, respectively. Elements of $GF(p^n)$ will usually be denoted by c, c_i; elements of R (in other words, polynomials in x over the Galois field) by A, B, C, D, H, M, N, P, where P denotes a typical irreducible polynomial in x. The poly-

* The cyclotomic polynomial $F_m(x)$ is the polynomial (with leading coefficient = 1) whose roots are the primitive mth roots of unity.

nomial M is said to be *primary* if the coefficient of the highest power of x occurring in M is the unit element of $GF(p^n)$. Typical elements of $R[u]$ or $F[u]$ will be denoted by $f(u)$, $g(u)$, $h(u)$. The degree of M (for M in R) has the obvious meaning; the degree of $f(u)$ means the degree in u. If the coefficient of the highest power of u occurring in $f(u)$ is the unit element of $R(x, p^n)$, $f(u)$ is primary.

According to the formula (1.5), $\omega_M(u)$ is defined by means of the function $\psi(t)$. However as the present paper is concerned only with algebraic properties, we shall define $\omega_M(u)$ directly and show that all the properties of the polynomials follow readily from the new definition. One possibility is to take (1.6) as the definition, but it is perhaps more satisfactory to proceed somewhat differently.

For defining properties* we shall take

$$(2.1)\qquad \begin{cases} \omega_{M+N} = \omega_M + \omega_N, & \\ \omega_{cM} = c\omega_M & (c \text{ in } GF(p^n)), \\ \omega_{x^{k+1}} = (\omega_{x^k})^{p^n} - x\omega_{x^k} & (k = 0, 1, 2, \cdots), \\ \omega_1 = u, & \end{cases}$$

where M, N are arbitrary polynomials in R, and for brevity we write ω_M in place of $\omega_M(u)$. Then it is easy to see, to begin with, that for all M,

$$(2.2)\qquad \omega_{xM} = \omega_M{}^{p^n} - x\omega_M,$$

thus generalizing the third equation in (2.1). Again if in that equation we take $k=0$, we have

$$\omega_x = \omega_1{}^{p^n} - x\omega_1 = u^{p^n} - xu;$$

combining this with (2.2) we see that

$$(2.3)\qquad \omega_{xM} = \omega_x(\omega_M).$$

In this equation replace M by xM; then (2.3) becomes

$$\omega_{x^2M} = \omega_x(\omega_{xM}) = \omega_x\{\omega_x(\omega_M)\} = \omega_{x^2}(\omega_M),$$

since by the third equation in (2.1)

$$\omega_{x^2} = \omega_x{}^{p^n} - x\omega_x = \omega_x(\omega_x).$$

Continuing in this way we may show by an easy induction on k that

$$(2.4)\qquad \omega_{x^kM} = \omega_{x^k}(\omega_M),$$

* If the operator Ω is defined by $\Omega u = u^{p^n}$, then the third equation in (2.1) implies $\omega_x = (\Omega - x)u$, and it is easy to see that generally $\omega_M = M(\Omega - x)u$, where $M(\Omega - x)$ is the operator obtained by substituting $\Omega - x$ for x in M.

thus generalizing (2.2). If now we take (2.4) together with the first two equations in (2.1), we have at once

$$\omega_{MN} = \omega_M(\omega_N) = \omega_N(\omega_M), \tag{2.5}$$

for arbitrary M, N. Thus we see that (1.8) follows from the new definition (2.1). For the sequel this property is apparently fundamental.

It is now not difficult to derive the explicit formula (1.6) for the polynomial $\omega_M(u)$. Because of the linearity (with respect to t) of the polynomial $\psi_k(t)$ it is sufficient to prove (1.6) in the case $M=x^m$. The formula is clearly true for $m=0$. Assume it true up to and including the value m. Then by the third equation in (2.1),

$$\begin{aligned}\omega_{x^{m+1}} &= \left\{\sum_{j=0}^{m}\frac{(-1)^{m-j}}{F_j}\psi_j(x^m)u^{p^{nj}}\right\}^{p^n} - x\sum_{j=0}^{m}\frac{(-1)^{m-j}}{F_j}\psi_j(x^m)u^{p^{nj}} \\ &= (-1)^{m+1}x\psi_0(x^m)u + \sum_{j=1}^{m+1}\frac{(-1)^{m+1-j}}{F_j}\{x\psi_j(x^m) + [j]\psi_{j-1}^{p^n}(x^m)\}u^{p^{nj}}.\end{aligned} \tag{2.6}$$

But it is easily seen that (1.7) implies*

$$\psi_k(xt) = x\psi_k(t) + [k]\psi_{k-1}^{p^n}(t);$$

also $\psi_0(x^m)=x^m$ and $F_k=[k]F_{k-1}^{p^n}$; thus (2.6) becomes

$$\omega_{x^{m+1}} = \sum_{j=0}^{m+1}\frac{(-1)^{m+1-j}}{F_j}\psi_j(x^{m+1})u^{p^{nj}};$$

this completes the induction, and therefore establishes (1.6). It is also evident from the induction that the coefficient of $u^{p^{nj}}$ in (1.6) is integral, that is, $\psi_j(M)/F_j$ is a polynomial in x. Our results may be summed up in the following:

THEOREM 1. *The polynomial $\omega_M(u)$ defined by* (2.1) *for all M (where M is a polynomial in x with coefficients in $GF(p^n)$) satisfies the equation* (2.5). *The polynomial has the explicit expression* (1.6) *in which the coefficients of $u^{p^{nj}}$ are polynomials in x. In particular $\omega_M(u)$ is linear in u† as well as in M.*

In the next place from (2.5) it follows that $\omega_M(u)$ is a divisor of $\omega_{MN}(u)$. The coefficients of u^i in the quotient are polynomials in x. This is a consequence of the following:

THEOREM 2. *If in the equation*

$$\begin{aligned}u^{m+r} &+ M_1u^{m+r-1} + \cdots + M_{m+r} \\ &= (u^m + A_1u^{m-1} + \cdots + A_m)(u^r + B_1u^{r-1} + \cdots + B_r)\end{aligned}$$

* Duke Mathematical Journal, vol. 1 (1935), p. 141.

† That is, of the form $\sum_i \alpha_i u^{p^{ni}}$.

all the coefficients M, A, B are rational functions of x in the $GF(p^n)$, then the M's are polynomials if and only if all the A's and B's are polynomials.

This is an analogue of a well-known theorem of Gauss; the proof need not be given.

Consider now two polynomials $\omega_M(u)$, $\omega_N(u)$. We seek the greatest common divisor (ω_M, ω_N). Clearly if A is a common divisor of M and N, then ω_A is a common divisor of ω_M and ω_N. Let $D=(N, M)$ the greatest common divisor of M and N—to make it unique assume D primary—then

$$D = AM + BN,$$

for properly chosen polynomials A, B. Then by the first of (2.1),

$$\omega_D = \omega_{AM} + \omega_{BN},$$

from which follows

$$\omega_D = f(u)\omega_M + g(u)\omega_N, \tag{2.7}$$

where $f(u)$ and $g(u)$ are polynomials in u (whose coefficients are polynomials in x). But (2.7) shows that any common divisor of ω_M and ω_N is necessarily a divisor of ω_D. This proves the following theorem:

THEOREM 3. *For arbitrary M, N, the greatest common divisor of ω_M and ω_N is determined by*

$$\omega_D = (\omega_M, \omega_N), \tag{2.8}$$

where $D=(M, N)$, the greatest common divisor of M and N.

If P is an irreducible polynomial in x, then it follows from (1.6) that

$$\omega_P(u) \equiv u^{p^{nk}} \pmod{P}, \tag{2.9}$$

where k is the degree of P. Therefore by (2.5),

$$\omega_{PM} \equiv (\omega_M)^{p^{nk}} \pmod{P}.$$

If then M and e are arbitrary, we have

$$\omega_{P^eM} \equiv (\omega_M)^{p^{nke}} \pmod{P}. \tag{2.10}$$

It will be convenient for a later purpose to alter slightly the notation for the greatest common divisor in order to indicate that we are reducing coefficients (mod P). We shall use the symbol $(f(u), g(u))_P$ to denote the G.C.D. in this situation. Thus usually,

$$(f(u), g(u))_P \not\equiv (f(u), g(u)) \pmod{P}.$$

In the present case, because of (2.7), the two symbols are equivalent, and we may state

THEOREM 4. *If MN is not a multiple of the irreducible polynomial P, then for $D=(M, N)$, we have*

$$(\omega_{P^eM}, \omega_{P^fN})_P \equiv (\omega_D)^{p^{nke}} \pmod{P}.$$

where P is of degree k, and $e \leqq f$.

The theorem is an immediate consequence of (2.8) and (2.10).

Finally, we ask whether ω_M can have repeated factors. Since by (1.6) the derivative with respect to u is $(-1)^m M$, which is independent of u, there can clearly be no repeated factor. Also if the polynomial be taken (mod P), this indicates that for $P \nmid M$ there is no repeated factor. If $P \mid M$, we make use of (2.10).

THEOREM 5. *The polynomial $\omega_M(u)$ has only simple factors in $F[u]$. For $P \nmid M$, $\omega_M(u)$ has no repeated factors* (mod P). *For $M=P^eN$, $P \nmid N$, $\omega_M(u) \equiv \omega_N{}^{p^{nke}}$* (mod P), *where ω_N has only simple factors.*

3. **Definition* of** $W(u)$. For convenience assume M primary. Then suppose $\omega_M(u)$ exhibited as a product of (necessarily distinct) primary polynomials f_i in $R[u]$:

$$\omega_M(u) = f_1(u)f_2(u) \cdots f_k(u), \tag{3.1}$$

so that the coefficients of $f_i(u)$ are polynomials in x. Consider those $f_i(u)$ in the right member of (3.1) that divide no $\omega_A(u)$, where A is a *proper* divisor of M. The product of those $f_i(u)$ is by definition $W_M(u)$, so that in particular $W_M(u)$ is primary. Since $A \mid M$ implies $\omega_A(u) \mid \omega_M(u)$, it is evident that $W_A(u) \mid \omega_M(u)$. Again for A a proper divisor of M, it is clear from the definition that $(W_M(u), W_A(u))=1$; thus $\omega_M(u)$ is divisible by the product $\prod W_A(u)$, extended over all A dividing M. On the other hand since to each $f_i(u)$ in the right member of (3.1) corresponds by the definition a unique $W_A(u)$ of which it is a divisor, it follows that

$$\omega_M(u) = \prod_{A \mid M} W_A(u). \tag{3.2}$$

By inversion we have for $W_M(u)$ the formula

$$W_M(u) = \prod_{M=AB} \{\omega_A(u)\}^{\mu(B)}, \tag{3.3}$$

where $\mu(B)$ is the Möbius function† for polynomials in $R(x, p^n)$.

From (3.3) certain properties of $W_M(u)$ are immediate. For example the degree of $W_M(u)$ is

---

* Cf. Kronecker, *Vorlesungen über Zahlentheorie*, vol. 1, 1901, p. 283.

† See American Journal of Mathematics, vol. 54 (1932), p. 39.

$$\phi(M) = \sum_{M=AB} \mu(A)\,|B| = |M| \prod_{P|M} (1 - |P|^{-1}), \tag{3.4}$$

the product extending over all irreducible divisors of M. Here $|M| = p^{nm}$, where m is the degree of M, so that $|M|$ is the degree of $\omega_M(u)$. Comparison of the degree of both members of (3.2) leads to

$$\sum_{A|M} \phi(A) = |M|,$$

which is of course a direct consequence of (3.4). We remark that $\phi(M)$ may be defined independently as the number of quantities in a reduced residue system (mod M).

In the next place we evaluate $W_M(0)$. Since $W_1(u) = \omega_1(u) = u$, $W_1(0) = 0$. We now assume $M \neq 1$. For $M = P^e$, P irreducible, it follows from (1.6) and (1.7) that $W_P(0) = \pm P$. In general (3.3) implies

$$W_M(u) = \prod_{M=AB} \left\{\frac{\omega_A(u)}{u}\right\}^{\mu(B)},$$

so that

$$W_M(0) = \prod_{M=AB} (-1)^a A^{\mu(B)} = \begin{cases} (-1)^k P & \text{for } M = P^e, \\ 1 & \text{otherwise,} \end{cases} \tag{3.5}$$

where a, k is the degree of A, P, respectively.

Suppose next that M and N are arbitrary, $M \neq N$; then (3.2) implies that $W_M(u)W_N(u)$ is a divisor of $\omega_{MN}(u)$. Since $\omega_A(u)$ has no repeated factors it follows at once that

$$(W_M, W_N) = 1 \qquad (M \neq N). \tag{3.6}$$

If the irreducible $P \nmid MN$, we may assert slightly more:

$$(W_M, W_N)_P \equiv 1 \qquad (M \neq N,\ P \nmid MN). \tag{3.7}$$

In the general case, note that for $P \nmid M$, (3.3) implies

$$W_{P^eM} = \prod_{M=AB} \{\omega_{P^eA}\}^{\mu(B)} \prod_{M=AB} \{\omega_{P^{e-1}A}\}^{\mu(PB)} \equiv \prod_{M=AB} \{\omega_A^{\mu(B)}\}^{p^e - p^{e-1}} \pmod{P},$$

by (2.10); and therefore (for $P \nmid M$)

$$W_{P^eM} \equiv W_M^{p^e - p^{e-1}} \pmod{P}. \tag{3.8}$$

Hence we conclude that for $P \nmid MN$

$$(W_{P^eM}, W_{P^fN})_P \equiv 1 \qquad (M \neq N), \tag{3.9}$$

while

$$(3.10) \qquad (W_{P^eM}, W_{P^fM})_P \equiv (W_M)^{p^e - p^{e-1}} \pmod{P}$$

for $e \leqq f$.

Consider* now the greatest common divisor of $W_M(\omega_N(u))$ and $W_N(\omega_M(u))$, or briefly

$$(W_M(\omega_N), W_N(\omega_M)).$$

We take first the case $(M, N)=1$. Then by (3.3) and (2.5),

$$(3.11) \qquad W_M(\omega_N) = \prod_{M=AB} (\omega_{AN})^{\mu(B)} = \prod_{M=AB} \prod_{AN=DE} W_D^{\mu(B)}.$$

Now since $(M, N)=1$, the factorization DE may be obtained by factoring A and N independently and then combining in all possible ways. Thus (3.11) becomes

$$(3.12) \qquad \prod_{M=ABC} \prod_{N=DE} W_{AD}^{\mu(B)} = \prod_{N=DE} \prod_{M=AH} (W_{AD})^{\sum_{H=BC}\mu(B)},$$

but

$$\sum_{B|H} \mu(B) = \begin{cases} 1 & \text{for } H = 1, \\ 0 & \text{otherwise}, \end{cases}$$

so that $M=AG$ reduces to $M=A$. Therefore by (3.10) and (3.12) we have

$$(3.13) \qquad W_M(\omega_N) = \prod_{D|N} W_{DM} \quad \text{for } (M, N) = 1.$$

Interchanging M and N, (3.13) becomes

$$(3.14) \qquad W_N(\omega_M) = \prod_{A|M} W_{AN}.$$

By (3.6) the greatest common divisor of $W_M(\omega_N)$ and $W_N(\omega_M)$ may be found by picking out the equal terms in the right member of (3.13) and (3.14). But $AN=DM$ together with $(M, N)=1$ implies $N|D$, whence $D=N$ and $A=M$. Thus for $(M, N)=1$,

$$(3.15) \qquad (W_M(\omega_N), W_N(\omega_M)) = W_{MN}.$$

Suppose next that the irreducible $P|M$; then from (3.3) follows

$$W_M(\omega_{P^e}) = \prod_{M=AB} (\omega_{P^eA})^{\mu(B)} = W_{P^eM}.$$

More generally if every irreducible divisor of A is also a divisor of M, we have similarly

* For the proof compare Netto, Archiv der Mathematik und Physik, (3), vol. 4 (1902), pp. 65–67.

$$W_M(\omega_A) = W_{AM}. \tag{3.16}$$

If now $M=\prod P^e$, $N=\prod P^f$ are arbitrary, put

$$M = M_0M_1, \quad M_0 = \prod P^e \quad (P|N), \quad M_1 = \prod P^e \quad (P\nmid N),$$
$$N = N_0N_1, \quad N_0 = \prod P^f \quad (P|N), \quad N_1 = \prod P^f \quad (P\nmid N),$$

so that

$$(M_0, M_1) = (N_0, N_1) = (M_1, N_1) = 1,$$

while M_0, N_0 have precisely the same irreducible divisors. Then by (3.16),

$$W_M(\omega_N) = W_{M_0M_1}(\omega_{N_0N_1}) = W_{N_0M_0M_1}(\omega_{N_1}),$$

and this in turn, by (3.13), implies

$$W_M(\omega_N) = \prod_{D|N_1} W_{DN_0M_0M_1}.$$

Interchanging M and N in this equation, we have

$$W_N(\omega_M) = \prod_{A|M_1} W_{AM_0N_0N_1}.$$

As above the condition for common factors in the right members is $DN_0M_0M_1 = AM_0N_0N_1$, that is, $DM_1 = AN_1$, whence $D=N$ and $A=M$, and therefore (3.15) holds generally. The case $M=N$ is included, for by (3.16), $W_M(\omega_M) = W_{M^2}$.

THEOREM 6. *For arbitrary M, N, the greatest common divisor*

$$(W_M(\omega_N), W_N(\omega_M)) = W_{MN}.$$

4. **Irreducibility of $W_M(u)$.** Let β be a root of $W_M(u)=0$ in a properly chosen $F_1 \supset F(x, p^n)$. If $(M, A)=1$, the identity (3.11) implies

$$W_M\{\omega_A(\beta)\} = \prod_{D|A} W_{DM}(\beta) = 0,$$

so that $\omega_A(\beta)$ is also a root of $W_M(u)$. Assume $\omega_A(\beta)=\beta$. Then the polynomial $\omega_A(u)-u=\omega_{A-1}(u)$ has a root in common with $W_M(u)$, from which it follows that $A\equiv 1 \pmod M$. Similarly for $(M, A)=(M, B)=1$, $\omega_A(\beta)=\omega_B(\beta)$ implies $A\equiv B \pmod M$. Thus it is clear that if β is any root of $W_M(u)=0$, then the quantities $\omega_A(\beta)$, where A ranges over a reduced residue system $\pmod M$, are distinct roots of $W_M(u)=0$; by calculating the degree of $W_M(u)$ it is easily seen that the $\omega_A(\beta)$ furnish all the roots.

We shall now show* that $W_M(u)$ is irreducible in $F[u]$ or what amounts

* Cf. Weber, *Algebra*, 2d edition, vol. 1, 1898, pp. 596–600.

to the same thing (by Theorems 2) in $R[u]$. Assume the factorization

$$W_M(u) = f(u)g(u), \tag{4.1}$$

where $f(u)$ is irreducible in $R[u]$. Let β be a root of $f(u)=0$ (in a field $F_1 \supset F$). By the above paragraph, if we can show that $\omega_A(\beta)$ is also a root of $f(u)=0$ for all $(A, M)=1$, it will follow that $f(u)$ coincides with $W_M(u)$, and therefore that $W_M(u)$ is irreducible in $R[u]$. Clearly it suffices to show that $\omega_P(\beta)$ is a root of $f(u)=0$ for all irreducible P not dividing M. Assume therefore that $f(\omega_P(\beta)) \neq 0$, so that necessarily $g(\omega_P(\beta))=0$. Thus we see that the polynomial $g(\omega_P(u))$ has a root in common with the irreducible $f(u)$, and therefore

$$f(u) \mid g(\omega_P(u)). \tag{4.2}$$

On the other hand, by (2.9), for P of degree k,

$$\omega_P(u) \equiv u^{p^{nk}} \pmod{P},$$

which implies

$$g(\omega_P(u)) \equiv g(u^{p^{nk}}) \equiv g^{p^{nk}}(u) \pmod{P}.$$

Comparison with (4.2) shows that

$$(f(u), g(u))_P \equiv h(u) \pmod{P},$$

where $h(u)$ is of positive degree in u. Thus (4.1) implies that $W_M(u)$ has a repeated factor (mod P); since $P \nmid M$, this contradicts Theorem 5. We may state the following:

THEOREM 7. *For arbitrary M, the polynomial $W_M(u)$ is irreducible in $F[u]$, the ring of polynomials in u with coefficients in the field $F(x, p^n)$ of rational functions of x in $GF(p^n)$.*

This theorem may be extended somewhat.* For $(M, N)=1$, let β be a root of $W_M(u)=0$, γ a root of $W_N(u)=0$. Then

$$\omega_{MN}(\beta+\gamma) = \omega_{MN}(\beta) + \omega_{MN}(\gamma) = \omega_N(\omega_M(\beta)) + \omega_M(\omega_N(\gamma)) = 0,$$

so that $\beta+\gamma$ is a root of $\omega_{MN}(u)=0$; indeed we shall now show that it is a root of $W_{MN}(u)=0$. For assume $W_D(\beta+\gamma)=0$, where D is a proper divisor of MN, from which follows $\omega_D(\beta+\gamma)=0$. Now $D=AB$, where $A \mid M$, $B \mid N$; we may suppose that A is a proper divisor of M. Then by (2.5), $\omega_{AB}(\beta+\gamma)=0$ implies $\omega_{AN}(\beta+\gamma)=0$; but as above

$$\omega_{AN}(\beta+\gamma) = \omega_A(\omega_N(\beta)) + \omega_A(\omega_N(\gamma)) = \omega_A(\omega_N(\beta)).$$

* Cf. Weber, loc. cit., pp. 600–601.

Since $(N, M)=1$, $\omega_N(\beta)$ is a root of $W_M(u)=0$ and therefore not of $\omega_A(u)=0$. This proves

$$W_{MN}(\beta+\gamma) = 0 \qquad (\text{for } (M, N) = 1). \tag{4.3}$$

Under the hypothesis $(M, N)=1$ we may choose A, B such that $AM+BN=1$. Put $\alpha=\beta+\gamma$, then

$$\begin{aligned}\omega_{AM}(\alpha) &= \omega_{AM}(\beta) + \omega_{AM}(\gamma)\\ &= \omega_{AM}(\beta) + \gamma - \omega_{BN}(\gamma)\\ &= \omega_A(\omega_M(\beta)) + \gamma - \omega_B(\omega_N(\gamma)),\end{aligned}$$

so that we have

$$\omega_{AM}(\alpha) = \gamma, \qquad \omega_{BN}(\alpha) = \beta. \tag{4.4}$$

Let us now assume that $W_M(u)$ factors in $F_1[u]$, where $F_1=F(\gamma)$ is the field obtained by adjoining γ to F: put

$$W_M(u) = f(u, \gamma)g(u, \gamma),$$

where $f(u, v)$, $g(u, v)$ are polynomials with coefficients in F. Let β be a root of $f(u, \gamma)=0$, then by (4.4)

$$f\{\omega_{BN}(\alpha), \omega_{AM}(\alpha)\} = 0.$$

But since $W_{MN}(u)$ is irreducible in $F[u]$, it follows from the first paragraph in this section that

$$f\{\omega_{BN}(\omega_H(\alpha)), \omega_{AM}(\omega_H(\alpha))\} = 0 \tag{4.5}$$

for all $(H, MN)=1$. Now for arbitrary $(D, M)=1$, we may choose H so that

$$H \equiv D \pmod{M}, \qquad H \equiv 1 \pmod{N}.$$

Since by (4.4),

$$\begin{aligned}\omega_{BN}(\omega_H(\alpha)) &= \omega_H(\omega_{BN}(\alpha)) = \omega_H(\beta) = \omega_D(\beta),\\ \omega_{AM}(\omega_H(\alpha)) &= \omega_H(\omega_{AM}(\alpha)) = \omega_H(\gamma) = \gamma,\end{aligned}$$

we have after substitution in (4.5),

$$f\{\omega_D(\beta), \gamma\} = 0,$$

so that $f(u, \gamma)$ has all the roots of $W_M(u)=0$.

THEOREM 8. *Let $(M, N)=1$, γ a root of $W_N(u)=0$, $F_1=F(\gamma)$ the field obtained by adjoining γ to F; then the polynomial $W_M(u)$ is irreducible in $F_1[u]$.*

As an application of Theorem 7 we state the following theorems:

THEOREM 9. *The group for the field F of the equation $W_M(u)=0$ is abelian; indeed it is simply isomorphic with the group (with respect to multiplication) of the reduced residue system* (mod M).

THEOREM 10. *If β is a root of $W_M(u)$, and t is an indeterminate, then the group for the field $F(\beta, t)$ of the equation $W_M(u)=t$ is abelian; indeed it is simply isomorphic with the additive group of residues* (mod M).

5. **Irreducibility proofs for the case $M=P^e$.** In the case $M=P^e$, where P, is irreducible, Theorem 7 may be proved very quickly in the following way. Assume the factorization

$$W_{P^e}(u) = f(u)g(u),$$

where $f(u)$ and $g(u)$ are in $R[u]$. By (3.5) $W_{P^e}(0)=\pm P$, so that $f(0)g(0)=\pm P$. We may therefore suppose that $f(0)=c$, an element of $GF(p^n)$. Construct the polynomial

$$h(u) = \prod_A f(\omega_A(u)), \tag{5.1}$$

where A ranges over a reduced residue system (mod P^e). Let β be an arbitrary root of $f(u)=0$. For fixed $A, P\nmid A$, determine B such that $AB=1+DP^e$. Thus

$$\omega_B(\omega_A(\beta)) = \omega_{BA}(\beta) = \beta + \omega_{DP^e}(\beta) + \beta + \omega_D(\omega_{P^e}(\beta)) = \beta,$$

$$f[\omega_B\{\omega_A(\beta)\}] = f(\beta) = 0,$$

from which it follows that

$$h(\omega_A(\beta)) = 0,$$

so that $h(u)=0$ is satisfied by every root of $W_M(u)=0$. Therefore $W_M(u)\mid h(u)$, and $W_M(0)\mid h(0)$. But $W_{P^e}(0)=\pm P$ and from (5.1) it follows at once that $h(0)=1$. This evidently proves our theorem.

It is clear from (1.6) that except for the coefficient of the highest power of u, all coefficients of $W_P(u)$ are divisible by P, while the last coefficient is $\pm P$. Let k be the degree of P; then by (2.5) and the last sentence, we have

$$W_{P^2}(u) = \frac{\omega_{P^2}(u)}{\omega_P(u)} = \frac{\omega_P(\omega_P(u))}{\omega_P(u)} = \{\omega_P(u)\}^{p^{nk}-1} + P\cdot g(u),$$

so that except for the leading term every coefficient is a multiple of P. The last term (that is, the one free of u) is precisely $(-1)^kP$. Clearly we may continue in this way and prove that in $W_{P^e}(u)$ every coefficient after the first is divisible by P, while by (3.5) the last term is $(-1)^kP$. Then the irreducibility of $W_{P^e}(u)$ in $R[u]$ follows as a special case of the following theorem:

THEOREM 11. *If in* $f(u)=u^k+A_1u^{k-1}+\cdots+A_k$, *all the* A_j *are divisible by some irreducible* P, *while* $P^2\nmid A_k$, *then* $f(u)$ *is irreducible in* $R[u]$.

Clearly this is an analogue of Eisenstein's well-known criterion for irreducibility. To prove the theorem, assume the factorization

$$f(u) = (u^r + M_1u^{r-1} + \cdots + M_r)(u^s + N_1u^{s-1} + \cdots + N_s).$$

We may suppose that $P \mid M_r$ while $P\nmid N_s$. Then

$$A_{k-1} = N_sM_{r-1} + N_{s-1}M_r;$$

since $P\mid A_{k-1}$, $P\mid M_r$, $P\nmid N_s$, it follows that $P\mid M_{r-1}$. Similarly from

$$A_{k-2} = N_sM_{r-2} + N_{s-1}M_{r-1} + N_{s-2}M_r,$$

it follows that $P\mid M_{r-2}$. Thus we prove that all the M_j are divisible by P. Consider now the coefficient of u^r:

$$A_s = N_s + N_{s-1}M_1 + N_{s-2}M_2 + \cdots.$$

Since $P\mid A_s$ this equation is certainly impossible. Hence $f(u)$ is irreducible.

6. **Factorization of W_M (mod P).** Assume first that $P\nmid M$. As usual let k be the degree of P. Let $e>0$ be the smallest exponent such that

$$P^e \equiv 1 \pmod{M}. \tag{6.1}$$

Then $e\mid\phi(M)$, where $\phi(M)$ is the Euler function for polynomials in R, and is evaluated by (3.4). We recall for later use that $\phi(M)$ is the degree of $W_M(u)$.

To begin with, we have

$$\omega_M(u) = \prod_{A\mid M} W_A(u); \tag{6.2}$$

secondly since $M\mid(P^e-1)$ it follows that

$$\omega_M(u) \mid \omega_{P^e-1}(u). \tag{6.3}$$

Next by (2.1) and (2.10),

$$\omega_{P^e-1}(u) = \omega_{P^e}(u) - \omega_1(u) \equiv u^{p^{nke}} - u \pmod{P}. \tag{6.4}$$

Now since P is irreducible, the complete set of residues (mod P) form a finite field, which is indeed a concrete representation of the $GF(p^{nk})$. Then by a well-known theorem, we have the identity

$$u^{p^{nke}} - u \equiv \prod_{\deg f\mid e} f(u) \pmod{P}, \tag{6.5}$$

the product extending over all $f(u)$ irreducible (mod P) and of degree a divisor of e. Then by (6.2), (6.3), (6.4), (6.5), it follows that $W_M(u)$ is con-

gruent (mod P) to the product of a certain number of the $f(u)$ occurring in the right member of (6.5). We shall now prove that in this factorization a polynomial $f(u)$ of degree $<e$ cannot appear. For suppose

$$f(u) \mid W_M(u) \pmod{P}, \tag{6.6}$$

where $f(u)$ is of degree $s<e$. By (6.5) we have also

$$f(u) \mid u^{p^{nke}} - u \pmod{P}. \tag{6.7}$$

Then (6.6) and (6.7) together imply

$$f(u) \mid (W_M(u),\ \omega_{P^e-1}(u))_P \pmod{P}, \tag{6.8}$$

since by (2.10)

$$u^{p^{nke}} - u \equiv \omega_{P^e-1}(u) \pmod{P}.$$

Using (3.2), it follows from (6.8) that for some $A \mid (P^e-1)$

$$f(u) \mid (W_M(u),\ W_A(u)) \pmod{P}.$$

Since e is the smallest exponent for which (6.1) holds, $M \neq A$; and since $P \nmid MA$, we have a contradiction with (3.7). Therefore we conclude that all the irreducible divisors (mod P) of $W_M(u)$ are of degree e. Comparing with the degree of $W_M(u)$, we have the following:

THEOREM 12. *For irreducible $P \nmid M$, let $e>0$ be the least exponent for which* $P^e \equiv 1$ (mod M). *Then*

$$W_M(u) \equiv f_1(u)f_2(u) \cdots f_r(u) \pmod{P}, \tag{6.9}$$

where the $f_i(u)$ are irreducible (mod P) *of degree e, and $er=\phi(M)$, as defined by* (3.4).

To remove the restriction on P we use (3.8). Then we have the more general theorem:

THEOREM 13. *For irreducible P, let $M=P^sM_1$, where $P \nmid M$. Let $e>0$ be the least exponent for which $P^e \equiv 1$* (mod M_1). *Then*

$$W_M(u) \equiv \{f_1(u) \cdots f_r(u)\}^{p^s-p^{s-1}} \pmod{P}, \tag{6.10}$$

where the $f_i(u)$ are irreducible (mod P) *of degree e, and $er=\phi(M_1)$.*

As an application we consider certain congruences. We take first

$$W_M(u) \equiv 0 \pmod{P}, \tag{6.11}$$

where $P \nmid M$. Since solutions occur only when $W_M(u)$ has linear factors (mod P), it is clear that P must be $\equiv 1$ (mod M). In that case there are precisely $\phi(M)$ solutions; if β is a particular solution, all solutions are furnished by $\omega_A(\beta)$, where A ranges over a reduced residue system (mod M).

Next consider the congruence

$$\omega_M(u) \equiv 0 \pmod{P}, \tag{6.12}$$

where $P \nmid M$. Let $D=(M, P-1)$, so that (as in deriving (2.7)) $D=AM+B(P-1)$, for properly chosen A, B. Then by (2.1)

$$\begin{aligned} \omega_D(u) &= g(u)\omega_M(u) + h(u)\omega_{P-1}(u) \\ &\equiv g(u)\omega_M(u) + H(u)(u^{p^{nk}} - u) \pmod{P}. \end{aligned} \tag{6.13}$$

Thus all solutions of (6.12) are also solutions of $\omega_D(u)\equiv 0 \pmod{P}$. We may therefore suppose in (6.11) that $P\equiv 1 \pmod{M}$. In this case we may show that (6.12) has $|M|$ solutions (where as above $|M|=p^{nm}$, $m=\deg M$). Indeed if we put $P-1=MD$, (2.9) and (2.5) imply

$$u^{p^{nk}} - u \equiv \omega_D(\omega_M(u)) \pmod{P}, \tag{6.14}$$

so that $\omega_M(u)$ divides $u^{p^{nk}}-u \pmod{P}$, and therefore the congruence (6.12) has the maximum number of solutions. Again a solution of (6.11) is also a solution of (6.12). Let β be a solution of $W_M(u)\equiv 0$. Then for arbitrary A we have

$$\omega_M(\omega_A(\beta)) \equiv \omega_A(\omega_M(\beta)) \equiv 0 \pmod{P},$$

so that $\omega_A(\beta)$ is a solution of $\omega_M(u)\equiv 0$. Assume next that $\omega_A(\beta)\equiv\omega_B(\beta)$, whence $\omega_{A-B}(\beta)\equiv 0$. But this implies

$$\omega_{A-B}(\omega_H(\beta)) \equiv \omega_H(\omega_{A-B}(\beta)),$$

and therefore $W_M(u) \mid \omega_{A-B}(u)$, so that $M \mid A-B$. Thus the $|M|$ roots of (6.12) are furnished by $\omega_A(\beta)$, where A ranges over a *complete* residue system (mod M). It is clear from the above that the roots of (6.11) may be described as the *primitive* roots of (6.12).

If as above $P-1=MD$, (6.14) holds and we see that

$$u^{p^{nk}} - u \equiv \prod_{\delta} \{\omega_M(u) - \delta\} \pmod{P}, \tag{6.15}$$

where δ ranges over the roots of $\omega_D(u)\equiv 0$. Since $u^{p^{nk}}-u$ is completely factorable (mod P) it follows that for fixed δ, the congruence

$$\omega_M(u) \equiv \delta \pmod{P} \tag{6.16}$$

has $|M|$ roots. If u_0 is a particular solution of (6.16), then $u_0+\mu$ is also a solution of (6.16), where μ is any solution of the congruence $\omega_M(u)\equiv 0$ (mod P). Clearly if δ is not a root of $\omega_D(u)\equiv 0$, the congruence (6.16) has no solutions. This follows from

$$\omega_D\{\omega_M(u) - \delta\} \equiv \omega_{P-1}(u) - \omega_D(\delta) \equiv u^{p^{nk}} - u - \omega_D(\delta) \equiv -\omega_D(\delta),$$

for all u (mod P). We may now state the following two theorems.*

THEOREM 14. *The congruence* (6.12) *is completely solvable if and only if* $P\equiv 1$ (mod M) *similarly for* (6.11). *If β is any root of* (6.11), *the general solution of* (6.11) *is* $\omega_A(\beta)$, *where A ranges over a reduced residue system* (mod M); *the general solution of* (6.12) *is* $\omega_B(\beta)$ *where B ranges over a complete residue system* (mod M).

THEOREM 15. *Let* $P-1=MD$. *The congruence* (6.16) *is solvable if and only if δ is a root of* $\omega_D(\delta)\equiv 0$ (mod P). *If u_0 is a particular solution of* (6.16), *then the general solution is furnished by* $u_0+\mu$, *where μ ranges over the roots of* $\omega_M(\mu)\equiv 0$ (mod P).

Finally we generalize the last theorem by removing the restriction $M\mid P-1$. Let $(P-1, M)=H$, so that $M=AH$, $P-1=BH$. Then if (6.16) is assumed solvable, we have

$$\omega_B(\delta) \equiv \omega_{BHA}(u) \equiv \omega_{P-1}(\omega_A(u)) \equiv 0,$$

so that a necessary condition is

$$\omega_B(\delta) \equiv 0 \pmod{P}. \tag{6.17}$$

Again for $A_1M+B_1(P-1)=H$, it follows readily that

$$\omega_H(u) \equiv \omega_{A_1}(\delta) \pmod{P}. \tag{6.18}$$

By Theorem 15, (6.17) is a sufficient condition for the solvability of (6.18). But if (6.18) holds, it is clear that

$$\omega_M(u) \equiv \omega_{AH}(u) \equiv \omega_{AA_1}(\delta) \equiv \omega_1(\delta) - \omega_{BB_1}(\delta) \equiv \delta,$$

so that (6.16) is indeed satisfied. Thus (6.17) is both necessary and sufficient for the solvability of (6.16). Also it is evident from the above that (6.16) has exactly the same solutions as (6.18). We have therefore the following:

THEOREM 16. *For arbitrary M, let* $(M, P-1)=H$, $M=AH$, $P-1=BH$, $AA_1+BB_1=1$. *Then the congruence* (6.16) *is solvable if and only if* (6.17) *holds; the congruences* (6.16) *and* (6.18) *are equivalent.*

THEOREM 17. *Let* $(M,P-1)=1$. *Then for arbitrary δ, the congruence* (6.16) *has a unique solution. Thus* (6.16) *defines a* (1, 1) *transformation of the residues* (mod P); *the inverse of the transformation is* $\omega_{A_1}(\delta)\equiv u$ (mod P), *where* $A_1M\equiv 1$ (mod $P-1$).

For in this case $B=P-1$, and (6.17) is automatically satisfied.

* Analogues of well known results on binomial congruences, modulo p.

DUKE UNIVERSITY,
DURHAM, N. C.

ANALYTICITY OF EQUILIBRIUM FIGURES OF ROTATION*

BY

BERNARD FRIEDMAN

Introduction

The problem of ascertaining the possible forms of relative equilibrium of a homogeneous gravitating mass of liquid, when rotating about a fixed axis with constant angular velocity, had its origin in the investigations on the theory of the earth's figure which began with Newton and MacLaurin. In recent times it has undergone much development especially at the hands of Poincaré, Liapounoff, and Lichtenstein.

We take the axis of rotation as the axis of z and the mass-center, which must evidently lie on the axis, as origin. If ω be the angular velocity of rotation, the component accelerations at (x, y, z) are $-\omega^2 x$, $-\omega^2 y$, $-\omega^2 z$ and the dynamical equations reduce to

$$-\omega^2 x = -\frac{1}{\rho}\frac{\partial p}{\partial x} - \frac{\partial \Omega}{\partial x}, \qquad -\omega^2 y = -\frac{1}{\rho}\frac{\partial p}{\partial y} - \frac{\partial \Omega}{\partial y}, \qquad 0 = -\frac{1}{\rho}\frac{\partial p}{\partial z} - \frac{\partial \Omega}{\partial z},$$

where Ω is the potential energy per unit mass, p the pressure, and ρ the density. Hence, integrating, we have

$$p/\rho = \tfrac{1}{2}\omega^2(x^2 + y^2) - \Omega + \text{const}.$$

At the free surface, p = constant and we have

$$\tfrac{1}{2}\omega^2(x^2 + y^2) - \int_R \frac{dV_Q}{\overline{PQ}} = \text{const.}, \tag{1}$$

where R is the region containing the rotating mass.

Liapounoff and Lichtenstein† have proved that, at all points where the apparent gravity is not zero, the surface possesses continuous derivatives of all orders but the problem of the analyticity has so far defied solution. This problem is equivalent in difficulty to that of the analyticity of the solutions of elliptic differential equations of the second order which was solved by E. Hopf.‡

---

* Presented to the Society, September 10, 1937; received by the editors July 6, 1936, and in revised form, January 11, 1937.

† Lichtenstein, *Gleichgewichtsfiguren rotierender Flussigkeiten.*

‡ Mathematische Zeitschrift, vol. 34 (1931), p. 194.

In this article I use the method developed by E. Hopf in the above paper to prove that the surface of equilibrium figures of rotation is analytic at all points where the apparent gravity exists, that is the gradient of the pressure is not zero.

The equation of the surface of revolution is given implicitly by equation (1). By a few simple transformations we generalize (1) so that it will have a meaning for complex values of x and y and then we differentiate partially with respect to x and y obtaining equations (10) of Part I.

Knowing that a solution exists for real values of x and y we set up a sequence of approximating non-monogenic functions (cf. equations (14), (15) of Part II) which reduce for x and y real to the known solution. We prove the sequence converges uniformly and that the limit is an analytic function.

I wish to express my thanks and gratitude to Professor E. Hopf who proposed the problem and without whose assistance and encouragement it would not have been solved.

I. Formulation of the Problem

We wish to prove that if R is any 3-dimensional region, B its boundary, which satisfies the following equation:

$$\int_R \frac{dV_Q}{\overline{PQ}} - F(P) = 0 \quad \text{for all } P \text{ on } B, \tag{1}$$

where $F(P)$ is an analytic function of P, dV_Q the element of volume and $\overline{PQ}$ the distance from P to Q, then the surface formed by B is analytic at all points P' where the gradient of (1) is not zero. It will be assumed that surfaces satisfying this equation possess a sufficient number of derivatives. This has been proved by Lichtenstein.*

Take any such point P' as $(0, 0, z_0)$, $z_0 > 0$, and let the z-axis be normal to B at P'. Let the equation of the surface be $z = z(x, y)$. Then since $z(x, y)$ has partial derivatives of all orders we have $z_x'(0, 0) = z_y'(0, 0) = 0$.

Since the gradient of (1) is not zero

$$\frac{\partial}{\partial z}\left[\int_R \frac{dV_Q}{\overline{P'Q}} - F(P')\right] \neq 0, \qquad P' = (0, 0, z_0).$$

Because of the continuity of the surface there then exist positive numbers r, r_2 such that for $x^2+y^2 < r_1^2$ we have $z(x, y) > 0$, $|z(x, y) - z_0| < r_2$ and

$$\frac{\partial}{\partial z}\left[\int_R \frac{dV_Q}{\overline{PQ}} - F(P)\right] > 4d \quad \text{for} \quad P = (x, y, z), \tag{2}$$

* Loc. cit.

where d is some positive constant.

Let $a<r_1$. Denote the semi-cylinder $\xi^2+\eta^2<a^2, \zeta>0$ by (a). Then (1) can be written as

$$(3)\qquad \int_{(a)\cdot R}\frac{dV_Q}{PQ}+\int_{R-(a)\cdot R}\frac{dV_Q}{PQ}-F(P)=0,\qquad Q=(\xi,\eta,\zeta).$$

The second integral, call it $G_a(x, y, z)$, is the potential at P of the region $R-(a)\cdot R$. $G_a(x, y, z)$ is known to be an analytic function of x, y, z as long as P is not in $R-(a)\cdot R$ that is if $x^2+y^2<a^2$, $|z-z_0|<r_2$.

Consider*

$$\begin{aligned}\frac{\partial}{\partial z}\int_{(a)\cdot R}\frac{dV_Q}{PQ}\Big|_{z=z(x,y)}&=\frac{\partial}{\partial z}\iint_a d\xi d\eta\int_0^z\frac{d\zeta}{[(\xi-x)^2+(\eta-y)^2+(\zeta-z)^2]^{1/2}}\Big|_{z=z(x,y)}\\&=\iint_a\frac{d\xi d\eta}{[\rho^2+z(x,y)^2]^{1/2}},\end{aligned}$$

where a denotes the circle $\xi^2+\eta^2<a^2$, $\rho^2=(\xi-x)^2+(\eta-y)^2$ and so

$$\left|\frac{\partial}{\partial z}\int_{(a)\cdot R}\frac{dV_Q}{PQ}\right|<2\iint_a\frac{d\xi d\eta}{\rho}.$$

By Schmidt's inequality,† we have, however

$$(4)\qquad \iint_a\frac{d\xi d\eta}{\rho}\leqq 2\pi\left(\frac{\pi a^2}{\pi}\right)^{1/2}=2\pi a,$$

so that

$$\left|\frac{\partial}{\partial z}\int_{(a)\cdot R}\frac{dV_Q}{PQ}\right|<4\pi a.$$

Taking $4\pi a<2d$ and using (2) and (3) we have

$$\left|\frac{\partial}{\partial z}(G_a-F)\right|>2d\quad\text{for}\quad x^2+y^2<a^2,\qquad |z-z_0|<r_2.$$

We can now write (1) as follows:

* Note that z means the third independent coordinate of the set (x, y, z) but $z(x, y)$ refers to the equation of the surface as a function of x, y.

† Sternberg, *Potentialtheorie*, Sammlung Göschen, p. 99, equation (3).

$$\iint_a d\xi d\eta \int_0^{z(\xi,\eta)} \frac{d\zeta}{[\rho^2 + (\zeta - z(x, y))^2]^{1/2}} + G_a - F = 0. \tag{5}$$

The integral can be split up into

$$\iint_a d\xi d\eta \left[\int_0^{z(x,y)} + \int_{z(x,y)}^{z(\xi,\eta)}\right] \frac{d\zeta}{[\rho^2 + (\zeta - z(x, y))^2]^{1/2}}.$$

After we make the change of variable $\zeta' = z(x, y) - \zeta$ and integrate, it becomes

$$\iint_a d\xi d\eta \left[f\left(\frac{z(x, y)}{\rho}\right) + f\left(\frac{z(x, y) - z(\xi, \eta)}{\rho}\right)\right],$$

where $f(u) = \log (u + (u^2+1)^{1/2})$ is regular for $u = 0$.

Call the integral of the first term $g_a(x, y, z)$, so that

$$g_a(x, y, z) = \iint_a [\log [z + (z^2 + \rho^2)^{1/2}] - \log \rho] d\xi d\eta.$$

$g_a(x, y, z)$ is an analytic function of x, y, z for $|z - z_0| < r_2$, $x^2+y^2 < a^2$ because $z + (z^2+\rho^2)^{1/2} > 0$ for all ξ, η and

$$\iint_a \log \rho \, d\xi d\eta = -\frac{\pi}{3}(x^2 + y^2) + \frac{a^2}{2}(\log a - \tfrac{1}{2}).$$

Now by (4)

$$\frac{\partial}{\partial z} g_a(x, y, z) = \iint_a \frac{d\xi d\eta}{(z^2 + \rho^2)^{1/2}} \leqq \iint_a \frac{d\xi d\eta}{\rho} \leqq 2\pi a,$$

so that

$$\left|\frac{\partial}{\partial z} g_a(x, y, z)\right| < d.$$

Call $g_a + G_a - F = H_a(x, y, z)$. Then H_a is an analytic function of x, y, z for $x^2+y^2 < a^2$, $|z - z_0| < r_2$ and

$$\left|\frac{\partial H_a}{\partial z}(x, y, z)\right| > d. \tag{6}$$

Equation (5) can now be written

$$\iint_a f\left(\frac{z(x, y) - z(\xi, \eta)}{\rho}\right) d\xi d\eta + H_a(x, y, z) = 0. \tag{7}$$

To put (7) in a more easily handled form we differentiate it with respect to x and y. We have

$$z_x'(x, y)\frac{\partial H_a}{\partial z}(x, y, z(x, y)) + \iint_a f'\left(\frac{z(x, y) - z(\xi, \eta)}{\rho}\right)\left[\frac{z_x'(x, y)}{\rho} - \frac{x-\xi}{\rho^3}(z(x, y) - z(\xi, \eta))\right] d\xi d\eta = -\frac{\partial H_a}{\partial x} \tag{8}$$

and a similar equation for z_y'.

Let

$$\frac{\partial H_a}{\partial z} = L_a, \qquad \frac{\partial H_a}{\partial x} = M_a, \qquad \frac{\partial H_a}{\partial y} = N_a, \tag{9}$$
$$z_x'(x, y) = z_1(x, y), \qquad z_y'(x, y) = z_2(x, y).$$

Call the integral in (8) $F_1(x, y)$ and the corresponding integral in the equation for z_y', $F_2(x, y)$.

Then our equations become, omitting for convenience the subscript a,

$$\begin{aligned} z_1(x, y)L(x, y, z(x, y)) + F_1(x, y) &= -M(x, y, z(x, y)),\\ z_2(x, y)L(x, y, z(x, y)) + F_2(x, y) &= -N(x, y, z(x, y)), \end{aligned} \tag{10}$$

$$\begin{aligned} &z(x, y) - z(a\cos\phi, a\sin\phi) = \int_{(a\cos\phi,\, a\sin\phi)}^{(x,y)} z_1(x', y')dx' + z_2(x', y')dy',\\ &z_1(0, 0) = z_2(0, 0) = 0. \end{aligned} \tag{11}$$

We wish now to consider (10) for complex values of x and y. To do this we shall extend the meaning of our integrals so that it will take account of complex values of x and y. Then let $z^{(0)}(x, y)$, $z_1^{(0)}(x, y)$, $z_2^{(0)}(x, y)$ be any continuous functions of the complex variables x, y which reduce for real x and real y to $z(x, y)$, $z_1(x, y)$, $z_2(x, y)$ the solutions of (10). We then set up by the method of successive approximations $Z(x, y)$, $Z_1(x, y)$, $Z_2(x, y)$ which satisfy the extended form of (10) also for complex values of x, y and reduce, for real x and real y, to $z(x, y)$, $z_1(x, y)$, $z_2(x, y)$.

Let $x = x' + ix''$, $y = y' + iy''$. We shall restrict x and y to the region R_γ, where R_γ is defined as follows:

$$x'^2 + y'^2 < a^2, \qquad (x''^2 + y''^2)^{1/2} < \gamma[a - (x'^2 + y'^2)^{1/2}]. \tag{12}$$

Note that R_γ is convex, for if (x_1, y_1) and (x_2, y_2) are in R_γ so is $(x, y) = t(x_1, y_1) + (1-t)(x_2, y_2)$, since

$$\begin{aligned} (x''^2 + y''^2)^{1/2} + \gamma(x'^2 + y'^2)^{1/2} &\leqq t(x_1''^2 + y_1''^2)^{1/2} + (1-t)(x_2''^2 + y_2''^2)^{1/2}\\ &\quad + \gamma t(x_1'^2 + y_1'^2)^{1/2} + (1-t)\gamma(x_2'^2 + y_2'^2)^{1/2}\\ &\leqq \gamma t a + \gamma(1-t)a = \gamma a. \end{aligned}$$

We now extend the meaning of the integrals in (10). In

$$(13) \qquad Z(x, y) - Z(a \cos \phi, a \sin \phi) = \int Z_1(x', y')dx' + Z(x', y')dy'$$

let $x' = \tau'(x - a \cos \phi) + a \cos \phi$, $y' = \tau'(y - a \sin \phi) + a \sin \phi$, $0 \leqq \tau' < 1$. Then

$$(14) \quad Z(x, y) - Z(a \cos \phi, a \sin \phi) = \int_0^1 [(x - a \cos \phi)Z_1 + (y - a \sin \phi)Z_2]d\tau'$$

which has a meaning for complex x and y. In

$$Z(x, y) - Z(\xi, \eta) = \int_{\xi,\eta}^{x,y} Z_1(\alpha, \beta)d\alpha + Z_2(\alpha, \beta)d\beta$$

let

$$\alpha = \xi + \tau(x - \xi), \qquad \beta = \eta + \tau(y - \eta), \qquad 0 \leqq \tau \leqq 1.$$

Then

$$(15) \qquad Z(x, y) - Z(\xi, \eta) = \int_0^1 [Z_1(\alpha, \beta)(x - \xi) + Z_2(\alpha, \beta)(y - \eta)]d\tau.$$

In $F_1(x, y)$ let

$$(16) \qquad \begin{aligned} \xi &= x + t(a \cos \phi - x), \qquad & 0 \leqq t < 1, \\ \eta &= y + t(a \sin \phi - y), \qquad & 0 \leqq \phi < 2\pi. \end{aligned}$$

The substitution is legitimate since $\xi^2 + \eta^2 < a^2$. Then

$$\rho = t[(a \cos \phi - x)^2 + (a \sin \phi - y)^2]^{1/2} = t\Phi(\phi), \qquad \text{say}.$$

Using (15) and (16) we have

$$\begin{aligned} Z(x, y) - Z(\xi, \eta) &= t \int_0^1 [Z_1(\alpha, \beta)(a \cos \phi - x) + Z_2(\alpha, \beta)(a \sin \phi - y)]d\tau \\ &= tZ_4(x, y, t, \phi), \qquad \text{say}, \end{aligned}$$

and then

$$(17) \qquad \begin{aligned} F_1(x, y) = \int_0^{2\pi} \int_0^1 \left[\left(\frac{z_4}{\Phi}\right)^2 + 1\right]^{-1/2} &\left[\frac{z_1(x, y)}{\Phi} + \frac{z_4(a \cos \phi - x)}{\Phi^3}\right] adt \\ &\cdot (a - x \cos \phi - y \sin \phi)d\phi \\ \equiv F(x, y, z_1, z_4), \quad \text{say}. \end{aligned}$$

Similarly $F_2(x, y) \equiv F(x, y, z_2, z_4)$. Note that (14), (15), (18), and (19) have meaning for complex x and y. Our equations can now be written as follows:

$$Z_1(x, y)L(x, y, Z(x, y)) + F(x, y, Z_1(x, y), Z_4(x, y)) = -M(x, y, Z(x, y)),$$
$$Z_2(x, y)L(x, y, Z(x, y)) + F(x, y, Z_2(x, y), Z_4(x, y)) = -N(x, y, Z(x, y)),$$
$$Z(x, y) - Z(a\cos\phi, a\sin\phi)$$
$$= \int_0^1 [(x - a\cos\phi)Z_1(x', y') + (y - a\sin\phi)Z_2(x', y')]d\tau'.$$

These are three equations for the unknowns $Z_1(x, y)$, $Z_2(x, y)$, and $Z(x, y)$. Actually by substituting the value of Z from the third equation into the first two, we have two equations in the unknowns $Z_1(x, y)$ and $Z_2(x, y)$. These equations will be considered in the following region:

$$x,\ y,\ \text{in } R_\gamma,\ |z_1|,\ |z_2| < c,$$

where c is a constant to be determined later

We shall now prove that $H(x, y, Z(x, y))$ is analytic for x, y in R_γ and also obtain bounds for its first and second derivatives when a is small. Now in $H_a = g_a + G_a - F$, g_a and F are obviously analytic in R_γ. G_a will be analytic if it can be shown that $(x-\xi)^2+(y-\eta)^2+(z(x, y)-\zeta)^2 \neq 0$, where x, y is in R_γ and ξ, η, ζ is in $R-(a)\cdot R$. This amounts to showing that

$$(x - a\cos\phi)^2 + (y - a\sin\phi)^2 + (Z(x, y) - Z(a\cos\phi, a\sin\phi))^2 \neq 0.$$

Let $x'^2+y'^2=r^2$. Now

$$|(x - a\cos\phi)^2 + (y - a\sin\phi)^2| \geqq \text{real part } [(x - a\cos\phi)^2 + (y - a\sin\phi)^2]$$
$$\geqq (x' - a\cos\phi)^2 + (y' - a\sin\phi)^2 - \gamma^2(a - r)^2$$

because of (12). Also from (14) we have

$$|Z(x, y) - Z(a\cos\phi, a\sin\phi)| \leqq c|x - a\cos\phi| + c|y - a\sin\phi|$$

and

$$|Z(x, y) - Z(a\cos\phi, a\sin\phi)|^2 \leqq c^2\{[(x' - a\cos\phi)^2 + \gamma^2(a - r)^2]^{1/2}$$
$$+ [(y' - a\sin\phi)^2 + \gamma^2(a - r)^2]^{1/2}\}^2,$$

where c is max $[|Z_1(x, y)|, |Z_2(x, y)|]$ for x, y in R_γ; so that

$$|(x - a\cos\phi)^2 + (y - a\sin\phi)^2| > |Z(x, y) - Z(a\cos\phi, a\sin\phi)|^2$$

if

$$(x' - a\cos\phi)^2 + (y' - a\sin\phi)^2 - \gamma^2(a - r)^2$$
$$> c^2\{[(x' - a\cos\phi)^2 + \gamma^2(a - r)^2]^{1/2} + [(y' - a\sin\phi)^2 + \gamma^2(a - r)^2]^{1/2}\}^2.$$

But if $\gamma < \frac{1}{3}$, then

$$\frac{(x'-a\cos\phi)^2+(y'-a\sin\phi)^2+\gamma^2(a-r)^2+\gamma^2(a-r)^2-3\gamma^2(a-r)^2}{\{[(x'-a\cos\phi)^2+\gamma^2(a-r)^2]^{1/2}+[(y'-a\sin\phi)^2+\gamma^2(a-r)^2]^{1/2}\}^2}$$

$$\geqq \frac{1}{2}-\frac{3\gamma^2(a-r)^2}{\{[(x'-a\cos\phi)^2+\gamma^2(a-r)^2]^{1/2}+[(y'-a\sin\phi)^2+\gamma^2(a-r)^2]^{1/2}\}^2}\geqq\frac{1}{6}\,.$$

Therefore if $c^2<\frac{1}{6}$, G_a will be analytic and so will H_a, L_a, M_a, and N_a.

Using the above inequality we have

$$\begin{aligned}&|(x-a\cos\phi)^2+(y-a\sin\phi)^2+[Z(x,y)-Z(a\cos\phi,a\sin\phi)]^2|\\ &\geqq(\tfrac{1}{6}-c^2)\{[(x'-a\cos\phi)^2+\gamma^2(a-r)^2]^{1/2}+[(y'-a\sin\phi)^2+\gamma^2(a-r)^2]^{1/2}\}\\ &\geqq(\tfrac{1}{6}-c^2)[(x'-a\cos\phi)^2+(y'-a\sin\phi)^2].\end{aligned}$$

Similarly for $a^2<\xi^2+\eta^2<r_3^2$ we have

$$|(x-\xi)^2+(y-\eta)^2|\geqq(x'-\xi)^2+(y'-\eta)^2-\gamma^2(a-r)^2$$

and

$$\begin{aligned}|Z(x,y)-Z(\xi,\eta)|&<|Z(x,y)-Z(a\cos\phi,a\sin\phi)|\\ &\qquad+|Z(a\cos\phi,a\sin\phi)-Z(\xi,\eta)|\\ &\leqq c[|x-a\cos\phi|+|a\cos\phi-\xi|+|y-a\sin\phi|+|a\sin\phi-\eta|]\\ &\leqq 4c[|x-\xi|+|y-\eta|],\end{aligned}$$

and proceeding as above we obtain the similar equality:

$$\begin{aligned}&|(x-\xi)^2+(y-\eta)^2+(Z(x,y)-Z(\xi,\eta))^2|\\ &\qquad\geqq(\tfrac{1}{6}-16c^2)[(x'-\xi)^2+(y'-\eta)^2].\end{aligned}\tag{18}$$

To obtain bounds for L, M, N, L_z', M_z', N_z' we proceed as follows: We have*

$$\frac{\partial G}{\partial x}=\iint_{B-(a)\cdot B}\frac{\cos(n,x)}{\overline{PQ}}\,d\omega_Q$$

and similar expressions for the derivatives with respect to y and z where $B-(a)B$ is the boundary of $R-(a)R$ and $d\omega_Q$ is the surface element. Since $z(\xi,\eta)$ has derivatives of all orders there exists $r_4<r_3$ such that for $\xi^2+\eta^2<r_4^2$

$$\cos(n,z)>c_1>0,$$

where n is normal to B at ξ, η, $z(\xi,\eta)$. Let B' denote the part of the surface for which $\xi^2+\eta^2>r_4^2$. Then using (18)

* Ibid., p. 124, equation 35.

$$\left|\frac{\partial G}{\partial x}\right| \leqq \left|\iint_{B'} \frac{\cos(n, x)}{\overline{PQ}}\, d\omega_Q\right| + \frac{1}{c_1}\iint_{a^2<\xi^2+\eta^2<r_3^2} \left(\tfrac{1}{6} - 16c^2\right) \frac{d\omega_Q'}{[(\xi - x')^2 + (\eta - y')^2]^{1/2}},$$

where $d\omega' = d\omega \cos(n, z)$ is the projection of $d\omega$ on the (x, y)-plane.

The integral over B' is a constant independent of a, while the second integral is by Schmidt's inequality less than $2\pi r_3$. Therefore,

$$\left|\frac{\partial G}{\partial x}\right| < k_1, \qquad x, y \text{ in } R_\gamma \tag{19}$$

and similar inequalities exist for the derivatives with respect to y and Z, where k_1 is a constant independent of a.

For the second derivatives we have

$$\frac{\partial^2 G}{\partial z \partial x} = \iint_{B-(a)\cdot B} \cos(n, z) \frac{\partial\left(\frac{1}{r}\right)}{\partial x}\, d\omega_Q$$

and similar expressions for $\partial^2 G/\partial z\partial y$, $\partial^2 G/\partial z^2$. As before

$$\left|\frac{\partial^2 G}{\partial z \partial x}\right| \leqq \left|\iint_{B'} \cos(n, z) \frac{\partial\left(\frac{1}{r}\right)}{\partial x}\, d\omega_Q\right| + \frac{1}{c_1}\iint_{a^2<\xi^2+\eta^2<r_3^2} \left(\tfrac{1}{6} - 16c^2\right) \frac{d\omega_Q'}{(\xi - x')^2 + (\eta - y')^2}$$

since

$$\left|\frac{\partial\left(\frac{1}{r}\right)}{\partial x}\right| = \left|\frac{x - \xi}{r}\cdot\frac{1}{r^2}\right| \leqq \frac{\frac{1}{6} - 16c^2}{(\xi - x')^2 + (\eta - y')^2}$$

from (18). The first integral is again a constant independent of a. The second integral is of the order of $\log[(a\cos\phi - x')^2 + (a\sin\phi - y')^2]$. Therefore

$$\left|\frac{\partial^2 G}{\partial z \partial x}\right| \leqq k_2 \log\left[(a\cos\phi - x')^2 + (a\sin\phi - y')^2\right], \tag{20}$$

where k_2 is independent of a. A similar result holds for the second derivatives with respect to yz and z^2.

Before going any further we must find a better bound for M and N. When

$a=0$, H_a reduces to the left side of (1) and since the z-axis is normal to the surface at $(0, 0, z_0)$ we have

$$L_0(0, 0, z_0) = 4d, \qquad M_0(0, 0, z_0) = 0.$$

Using (7), (19), and Schmidt's inequality

$$| L_0(x, y, Z(x, y)) - L_a(x, y, Z(x, y)) | < k_3 a,$$

with similar inequalities for M and N. Since L_a, M_a, and N_a are analytic we can therefore choose a and γ so small that

$$d/2 < \quad | L_a |, \qquad | M_a |, \qquad | N_a | \quad < \frac{dc}{2} \quad \text{for} \quad x, y \text{ in } R_\gamma.$$

Since F and g_a are analytic they are bounded and using (20) we shall have

$$| L'_z | < k \log [(a \cos \phi - x')^2 + (a \sin \phi - y')^2],$$

and the same bound for M'_z and N'_z.

II. Proof of Analyticity

We have to consider the equations

$$\text{(1)} \qquad \begin{aligned} &Z_1(x, y)L(x, y, Z) + F(x, y, Z_1, Z_4) = - M(x, y, Z), \\ &Z_2(x, y)L(x, y, Z) + F(x, y, Z_2, Z_4) = - N(x, y, Z), \\ &Z(x, y) - Z(a \cos \phi, a \sin \phi) = \int_0^1 [(a \cos \phi - x)Z_1 + (a \sin \phi - y)Z_2] d\tau', \end{aligned}$$

where L, M, N are analytic functions of x, y, z for x, y in R_γ, $|z_1|$, $|z_2| < c$ and satisfy the following inequalities in that region:

$$\text{(2)} \qquad \begin{aligned} &k > | L | > d; \qquad | M |, | N | < \frac{dc}{2}; \\ &| N'_z |, \qquad | M'_z |, \qquad | L'_z | < k \log [(a \cos \phi - x')^2 + (a \sin \phi - y')^2]. \end{aligned}$$

F is defined as follows:

$$\text{(3)} \qquad F(x, y, Z_1, Z_4) = \int_0^{2\pi} \int_0^1 \frac{1}{\left(\left(\frac{Z_4}{\Phi}\right)^2 + 1\right)^{1/2}} \left[Z_1(x, y) + \frac{Z_4}{\Phi} \frac{a \cos \phi - x}{\Phi} \right] a \, dt \cdot \frac{a - x \cos \phi - y \sin \phi}{\Phi} d\phi,$$

where

$$\text{(4)} \qquad \Phi^2 = (a \cos \phi - x)^2 + (a \sin \phi - y)^2,$$

and

(5) $$Z_4(x, y, t, \phi) = \int_0^1 [Z_1(\alpha, \beta)(a \cos \phi - x) + Z_2(\alpha, \beta)(a \sin \phi - y)]d\tau,$$

where

(6) $$\alpha = \xi + \tau(x - \xi), \qquad \beta = \eta + \tau(y - \eta), \qquad 0 \leqq \tau \leqq 1$$

and

(7) $$\begin{aligned} \xi &= x + t(a \cos \phi - x), \qquad 0 \leqq t < 1, \\ \eta &= y + t(a \sin \phi - y), \qquad 0 \leqq \phi < 2\pi. \end{aligned}$$

If we let $x = x' + ix''$, $y = y' + iy''$ when x, y is in R_γ, that is,

(8) $$x'^2 + y'^2 < a^2, \qquad (x''^2 + y''^2)^{1/2} < \gamma[a - (x'^2 + y'^2)^{1/2}],$$

then $\Phi = 0$ if and only if $x = a \cos \phi$, $y = a \sin \phi$. For

$$\Phi^2 = [ae^{i\phi} - x - iy][ae^{-i\phi} - x + iy].$$

Now if $\Phi = 0$, one of the brackets is zero. Assume $ae^{i\phi} = x + iy$. Then

(9) $$| a | = | x + iy |$$

or

(10) $$a^2 = (x' - y'')^2 + (x'' - y')^2 = (x'^2 + y'^2) + x''^2 + y''^2 + 2(x''y' - x'y'').$$

By Cauchy's inequality,

$$| x''y' - x'y'' | \leqq (x'^2 + y'^2)^{1/2}(x''^2 + y''^2)^{1/2} \leqq \gamma a(x'^2 + y'^2)^{1/2} - \gamma(x'^2 + y'^2)$$

so that

$$\begin{aligned} (x' - y'')^2 + (x'' - y')^2 &\leqq x'^2 + y'^2 + \gamma^2[a - (x'^2 + y'^2)^{1/2}]^2 \\ &\quad + 2\gamma(x'^2 + y'^2)^{1/2}[a - (x'^2 + y'^2)^{1/2}] \\ &= [(x'^2 + y'^2)^{1/2} + \gamma(a - (x'^2 + y'^2)^{1/2})]^2 \\ &= [\gamma a + (1 - \gamma)(x'^2 + y'^2)^{1/2}]^2. \end{aligned}$$

Now this bracket is less than a^2 if $x'^2 + y'^2 < a^2$. So (10) can not be true and $\Phi \neq 0$. Hence for $\Phi = 0$ we must have $x'^2 + y'^2 = a^2$ which implies $x'' = y'' = 0$ and $x = a \cos \phi$, $y = a \sin \phi$. Therefore

(11) $$\left| \frac{a \cos \phi - x}{\Phi} \right|, \qquad \left| \frac{a \sin \phi - y}{\Phi} \right|$$

are bounded. By taking γ sufficiently small the upper bound of these can be made less than three so that

$$\text{(12)}\qquad \left|\frac{a\cos\phi - x}{\Phi}\right|, \qquad \left|\frac{a\sin\phi - y}{\Phi}\right| < 3 \quad \text{if} \quad x, y \text{ in } R_\gamma.$$

Then

$$\left|\frac{a - x\cos\phi - y\sin\phi}{\Phi}\right| = \left|\frac{\cos\phi(a\cos\phi - x) + \sin\phi(a\sin\phi - y)}{\Phi}\right| \leqq 5,$$

and calling $Z_4/\Phi = Z_5$ we have

$$\text{(13)}\qquad \begin{aligned} |Z_5| &\leqq \left|\frac{a\cos\phi - x}{\Phi}\right|\left|\int_0^1 |z_1|\,d\tau + \left|\frac{a\sin\phi - y}{\Phi}\right|\right|\int_0^1 |z_2|\,d\tau \\ &\leqq 3\int_0^1 [|z_1| + |z_2|]d\tau. \end{aligned}$$

We now define successive approximations $z^{(\nu)}, z_1^{(\nu)}, z_2^{(\nu)}$ to $Z(x, y), Z_1(x, y), Z_2(x, y)$ as follows:

$$\text{(14)}\qquad \begin{aligned} z_1^{(\nu+1)} &= -\frac{F(x, y, z_1^{(\nu)}, z_4^{(\nu)})}{L(x, y, z^{(\nu)})} - \frac{M(x, y, z^{(\nu)})}{L(x, y, z^{(\nu)})}, \\ z_2^{(\nu+1)} &= -\frac{F(x, y, z_2^{(\nu)}, z_4^{(\nu)})}{L(x, y, z^{(\nu)})} - \frac{N(x, y, z^{(\nu)})}{L(x, y, z^{(\nu)})}, \end{aligned}$$

$$\text{(15)}\ z^{(\nu)}(x, y) - z(a\cos\phi, a\sin\phi) = \int_0^1 [(a\cos\phi - x)z_1^{(\nu)} + (a\sin\phi - y)z_2^{(\nu)}]d\tau',$$

where

$$\text{(16)}\qquad z_4^{(\nu)} = \int_0^1 [z_1^{(\nu)}(\alpha, \beta)(a\cos\phi - x) + z_2^{(\nu)}(a\sin\phi - y)]d\tau.$$

The first approximations $z_1^{(0)}(x, y)$ and $z_2^{(0)}(x, y)$ are any continuous functions of the complex variables x, y which reduce, when x, y are real, to $z_1(x, y)$ and $z_2(x, y)$, the solutions of (1) in the real domain. By taking a and γ small enough we shall have

$$\text{(17)}\qquad |z_1^{(0)}|, \qquad |z_2^{(0)}| < c \quad \text{for} \quad x, y \text{ in } R_\gamma.$$

Then from (15) and (16) we have

$$\text{(18)}\ |z^{(0)}(x, y) - z(a\cos\phi, a\sin\phi)| < c[|a\cos\phi - x| + |a\sin\phi - y|] \leqq 3ac$$

and $|z_5{}^0| < 6c$.

Assume that

$$\text{(19)}\qquad |z_1^{(\nu)}|, \qquad |z_2^{(\nu)}| < c \quad \text{for} \quad x, y \text{ in } R_\gamma.$$

Then from (15) and (16) we have

(20) $$|z^{(\nu)}(x, y) - z(a \cos \phi, a \sin \phi)| < 3ac,$$

(21) $$|z_5^{(\nu)}| < 6c.$$

But from (14) and (2)

$$|z_1^{(\nu+1)}| < \frac{1}{d} F(x, y, z_1^{(\nu)}, z_4^{(\nu)}) + \frac{c}{2}.$$

Now

$$|F| \leqq \int_0^{2\pi} \int_0^1 [(z_5^{(\nu)})^2 + 1]^{-1/2} \left[z_1^{(\nu)} + z_5^{(\nu)} \frac{a \cos \phi - x}{\Phi} \right] dt$$
$$\cdot a \cdot \frac{a - x \cos \phi - y \sin \phi}{\Phi} d\phi \leqq 2\pi a(1 - 36c^2)^{-1/2}(5c + 90c).$$

Take $380\pi a < d(1-36c^2)^{1/2}$ so that

(22) $$|F|' < \tfrac{1}{2}cd,$$

and then $|z_1^{(\nu+1)}| < \frac{1}{2}c + \frac{1}{2}c = c$. Therefore (19) holds for all ν.

We define an operator Δ^ν as follows:

$$\Delta^\nu f(x, y, z_i) = f(x, y, z_i^{(\nu)}) - f(x, y, z_i^{(\nu-1)}) = \int_{z_i^{(\nu-1)}}^{z_i^{(\nu)}} \frac{\partial f}{\partial z_i} dz_i.$$

The integral is to be taken along the straight line joining $z_i^{(\nu-1)}$ to $z_i^{(\nu)}$. Notice that

(23) $$\Delta^\nu f(x, y, z_i)g(x, y, z_j) = g(x, y, z_j^{(\nu)})\Delta^\nu f(x, y, z_i) + f(x, y, z_i^{(\nu-1)})\Delta^\nu g(x, y, z_j).$$

By $z_i^{(\nu)}$ we mean the νth approximation to z_i.

Let

$$\max [|\Delta^\nu z_1|, |\Delta^\nu z_2|] = \sigma_\nu \quad \text{for} \quad x, y \text{ in } R.$$

From (14), using (23), we have

(24) $$\Delta^{\nu+1} z_1 = \Delta^\nu \frac{F}{L} + \Delta^\nu \frac{M}{L} = \frac{\Delta^\nu F}{L(x, y, z^{(\nu)})} + F(x, y, z_1^{(\nu-1)}, z_4^{(\nu-1)})\Delta^\nu \frac{1}{L}$$
$$+ \frac{\Delta^\nu M}{L(x, y, z^{(\nu)})} + M(x, y, z^{(\nu-1)})\Delta^\nu \frac{1}{L}\cdot$$

Now

(25) $$\Delta^\nu F = \iint [\Delta^\nu z_1(z_5^2 + 1)^{-1/2} + \Delta^\nu z_5(z_5^2 + 1)^{-1/2} \frac{a \cos \phi - x}{\Phi} \Big] a\, dt$$
$$\cdot \frac{a - x \cos \phi - y \sin \phi}{\Phi} d\phi.$$

But

$$|\Delta^{\nu} z_1(z_5^{\nu}+1)^{-1/2}| \leqq |z_1^{(\nu)}\Delta^{\nu}(z_5^2+1)^{-1/2}| + |[(z_5^{(\nu-1)})^2+1]^{-1/2}\Delta^{\nu} z_1|, \tag{26}$$

and using (21),

$$\begin{aligned} |\Delta^{\nu}(z_5^2+1)^{-1/2}| &= \left|\int_{z_5^{(\nu-1)}}^{z_5^{(\nu)}} z_5(z_5^2+1)^{-3/2}dz_5\right| \\ &\leqq 6c(1-36c^2)^{-3/2}|z_5^{(\nu)}-z_5^{(\nu-1)}|. \end{aligned} \tag{27}$$

Hence in (26)

$$\begin{aligned} |\Delta^{\nu} z_1(z_5^2+1)^{-1/2}| &< 6c^2(1-36c^2)^{-3/2}|z_5^{(\nu)}-z_5^{(\nu-1)}| \\ &\quad + (1-36c^2)^{-1/2}|\Delta^{\nu} z_1|. \end{aligned} \tag{28}$$

Considering the second term in (25), we have

$$\begin{aligned} |\Delta^{\nu} z_5(z_5^2+1)^{-1/2}| &= \left|\int_{z_5^{(\nu-1)}}^{z_5^{(\nu)}} (z_5^2+1)^{-3/2}dz_5\right| \\ &\leqq (1-36c^2)^{-3/2}|z_5^{(\nu)}-z_5^{(\nu-1)}|. \end{aligned} \tag{29}$$

From (16) we have, using (11),

$$|\Delta^{\nu} z_5| \leqq 6\sigma_\nu. \tag{30}$$

Hence in (25) using (12), (26), (28), (29), and (30), we have

$$|\Delta^{\nu} F| \leqq 2\pi a(1-36c^2)^{-1/2}[180c^2(1-36c^2)^{-1}+5+90]\sigma_\nu = ad_1\sigma_\nu, \quad \text{say}. \tag{31}$$

We also have

$$\begin{aligned} |\Delta^{\nu} L^{-1}| &\leqq \left|\int_{z^{(\nu-1)}}^{z^{(\nu)}} L_z' L^{-2}dz\right| \\ &\leqq kd^{-2}\log[(a\cos\phi-x')^2+(a\sin\phi-y')^2]\cdot|\Delta^{\nu} z|, \end{aligned} \tag{32}$$

$$\begin{aligned} |\Delta^{\nu} M| &\leqq \left|\int_{z^{(\nu-1)}}^{z^{(\nu)}} M_z' dz\right| \\ &\leqq k\log[(a\cos\phi-x')^2+(a\sin\phi-y')^2]\cdot|\Delta^{\nu} z|. \end{aligned} \tag{33}$$

But from (15)

$$|\Delta^{\nu} z| < \sigma_\nu[|a\cos\phi - x| + |a\sin\phi - y|], \tag{34}$$

and, using (31), (32), (33), (34), (21), and (2) in (24), we have

$$|\Delta^{\nu+1} z_1| < ad^{-1}(d_1 + \tfrac{1}{2}ck\log a + k\log a + k^2d^{-1}\log a)\sigma_\nu. \tag{35}$$

Taking a so small that

$$ad^{-1}(d_1 + \tfrac{1}{2}ck \log a + k \log a + k^2 d^{-1} \log a) < \tfrac{1}{2},$$

we have $|\Delta^{\nu+1} z_1| < \frac{1}{2}\sigma_\nu$ and a similar expression for $|\Delta^{\nu+1} z_2|$. Therefore $\sigma_{\nu+1} < \frac{1}{2}\sigma_\nu$ and σ_ν approaches zero since $\sigma_\nu < (\frac{1}{2})^\nu \sigma_0$. From (34) $|\Delta^\nu z|$ approaches zero and hence $z^{(\nu)}$, $z_1^{(\nu)}$, $z_2^{(\nu)}$ approach uniformly limit functions $Z(x, y)$, $Z_1(x, y)$, $Z_2(x, y)$.

We now wish to show that $Z(x, y)$ is an analytic function of x and y. Assume that $z_1^{(0)}(x, y)$ and $z_2^{(0)}(x, y)$ have continuous first partial derivatives with respect to x', x'' and y', y'', e.g., by assuming $Z_1^{(0)}(x, y) = z_1(x', y')$; $Z_2^{(0)}(x, y) = z_2(x', y')$. Then $z_1^{(0)}$ and $z_4^{(0)}$ will also have partial derivatives of the first order. Mathematical induction shows that the same is true for $z_1^{(\nu)}$, $z_2^{(\nu)}$, $z^{(\nu)}$, and $z_4^{(\nu)}$.

Consider the operators

$$\nabla_1 = \frac{\partial}{\partial x'} + i \frac{\partial}{\partial x''}, \qquad \nabla_2 = \frac{\partial}{\partial y'} + i \frac{\partial}{\partial y''}\cdot$$

Applying ∇_1 to (15) we have

$$\nabla_1 z_4^{(\nu)} = \int_0^1 [\nabla_1 z_1^{(\nu)}(a \cos \phi - x) + \nabla_1 z_2^{(\nu)}(a \sin \phi - y)] d\tau. \tag{36}$$

Applying it to (14), we have

$$\nabla_1 z^{(\nu)} = \int [\nabla_1 z_1^{(\nu)}(a \cos \phi - x) + \nabla_1 z_2^{(\nu)}(a \sin \phi - y)] d\tau', \tag{37}$$

since x and y are analytic. Also

$$\begin{aligned} \nabla_1 z_1^{(\nu+1)} &= -\frac{F(x, y, z_1^{(\nu)}, z_4^{(\nu)})}{L(x, y, z^{(\nu)})^2} L_z'(x, y, z^{(\nu)}) \nabla_1 z^{(\nu)} \\ &= -\frac{1}{L} \iint \Big\{ [(z_5^{(\nu)})^2 + 1]^{-1/2} \nabla_1 z_1^{(\nu)} - z_1^{(\nu)} z_5^{(\nu)} [(z_5^{(\nu)})^2 + 1]^{-3/2} \nabla_1 z_5^{(\nu)} \\ &\quad - \frac{a \cos \phi - x}{\Phi} [(z_5^{(\nu)})^2 + 1]^{-3/2} \nabla_1 z_5^{(\nu)} \Big\} a dt \\ &\quad \cdot \frac{a - x \cos \phi - y \sin \phi}{\Phi} d\phi - L^{-2}(LM_z' - L_z' M) \nabla_1 z^{(\nu)}. \end{aligned} \tag{38}$$

Let $\max [|\nabla_1 z^{(\nu)}|, |\nabla_1 z_2^{(\nu)}|] = \alpha_\nu$ for x, y in R_γ, then in (36) we have

$$|\nabla_1 z_5^{(\nu)}| < 6\alpha_\nu.$$

In (37) we have

$$|\nabla_1 z^{(\nu)}| < \alpha_\nu[|a\cos\phi - x| + |a\sin\phi - y|].$$

Using (21) and (2), we see that the first term on the right of (38) is less than $\frac{1}{2}ck\log[(a\cos\phi-x')^2+(a\sin\phi-y')^2]\cdot(|a\cos\phi-x|+|a\sin\phi-y|)$. Using (21), (11), and (12), we have that the integral in (38) is less than

$$a(1-36c^2)^{-1/2}[5d^{-1}+108ac^2(1-36c^2)^{-1}+90]\alpha_\nu = d_2a\alpha_\nu, \text{ say}.$$

The third term is less than $k^2d^{-2}a\log a\cdot\alpha_\nu$.

Hence in (38)

$$|\nabla_1 z_1^{(\nu+1)}| < [\tfrac{1}{2}ck\,a\log a + d_2a + k^2d^{-2}a\log a]\alpha_\nu.$$

Choose a so that

$$\tfrac{1}{2}ck\,a\log a + d_2a + k^2d^{-2}a\log a < \tfrac{1}{2}.$$

Then $|\nabla_1 z_1^{(\nu+1)}| < \frac{1}{2}\alpha_\nu$ and $\alpha_{\nu+1} < \frac{1}{2}\alpha_\nu < (\frac{1}{2})^{\nu+1}\alpha_0$ so that α_ν approaches zero. Since $|\nabla_1 z^{(\nu)}| < 4a\alpha_\nu$, $\nabla_1 z^{(\nu)}$ also approaches zero uniformly in x and y and therefore $\nabla_1 Z(x, y)$ exists and equals zero. This proves that $Z(x, y)$ is analytic in x. A similar proof holds for the analyticity in y.

The proof may seem irregular because we have varied our choice of a. But it can be seen that all we have required of it is that it satisfies the following inequalities:

$$380\pi a < d(1-36c^2)^{-1/2},$$
$$ad^{-1}(d_1 + \tfrac{1}{2}ck\log a + k\log a + k^2d^{-1}\log a) < \tfrac{1}{2},$$
$$\tfrac{1}{2}ck\,a\log a + d_2a + k^2d^{-2}a\log a < \tfrac{1}{2},$$

which can be done since the constants c, k, and d were proved to be independent of a.

Massachusetts Institute of Technology
Cambridge, Mass.

GENERALIZATIONS OF THE GAUSS LAW OF THE SPHERICAL MEAN*

BY

HILLEL PORITSKY

1. **Introduction.** The nature of the generalizations of the Gauss law of the spherical mean considered in this paper is illustrated by the following theorem (§3):

If in three-space the function u satisfies the differential equation

$$\nabla^2 u - \lambda u = 0, \qquad \lambda = \text{const.} \tag{1}$$

on and within a three-dimensional sphere S of radius r, then

$$A(u) = u_0 \sinh (r\lambda^{1/2})/(r\lambda^{1/2}), \tag{2}$$

where $A(u)$ is the average or arithmetic mean of u over S, that is

$$A(u) = \int u dS \Big/ \int dS, \tag{3}$$

while u_0 is the value of u at the center of S.

A similar result holds for any number of dimensions n. Thus for $n=2$, if u is a two-dimensional solution of (1),

$$A(u) = u_0 I_0(r\lambda^{1/2}), \tag{4}$$

where $A(u)$ is the mean of u over S, a circle of radius r. The general case of n dimensions is given by

$$A(u) = u_0\left[1 + \frac{r^2\lambda}{2\cdot n} + \frac{r^4\lambda^2}{2\cdot 4\cdot n(n+2)} + \cdots\right]. \tag{5}$$

We shall denote the bracket in (5) by $\phi_n(r, \lambda)$:

w.

2k o. $$\phi_n(r, \lambda) = \sum_{k=0}^{\infty} (r^2\lambda)^k/2\cdot 4 \cdots 2k\cdot n(n+2)\cdots(n+2k-2).$$

throı

(17), unction is expressible in terms of Bessel functions of order $n/2-1$ as s:

(20) $$\phi_n(r, \lambda) = \Gamma(n/2)(r\lambda^{1/2}/2)^{1-n/2} I_{n/1-2}(r\lambda^{1/2}).$$

* Presented to the Society, December 27, 1928; received by the editors March 9, 1937.

The familiar Gauss law of the mean for harmonic functions in n dimensions in obtained by putting $\lambda=0$ in (1) and (5).

Similar laws are derived below for solutions of

$$(\nabla^2 - \lambda)^p u = 0 \tag{8}$$

(§3); for these, it is shown that

$$\begin{aligned} A(u) &= u_0\phi_n(r, \lambda) + (\nabla^2 - \lambda)u]_0 \frac{\partial\phi_n(r, \lambda)}{\partial\lambda} + \cdots \\ &+ \frac{(\nabla^2 - \lambda)^{p-1}u]_0}{(p-1)!} \frac{\partial^{p-1}\phi_n(r, \lambda)}{\partial\lambda^{p-1}}, \end{aligned} \tag{9}$$

where the subscripts 0 following the brackets indicate evaluation at the center of the sphere. Alternative forms for the mean, say in terms of Bessel functions, are obtained from the relations

$$\begin{aligned} \frac{\partial^k\phi_n(r, \lambda)}{\partial\lambda^k} &= \Gamma(n/2)\lambda^{-k}(r\lambda^{1/2}/2)^{k-n/2+1}I_{n/2+k-1}(r\lambda^{1/2}) \\ &= [r^{2k}/2^k n(n+2)\cdots(n+2k-2)]\phi_{n+2k}(r, \lambda). \end{aligned} \tag{10}$$

Again the case $\lambda=0$ is of special interest; now (8) becomes the repeated Laplace equation:

$$\nabla^{2p}u = 0, \tag{11}$$

whose solutions are sometimes known as "p-harmonic" functions, while (9) reduces to

$$\begin{aligned} A(u) &= u_0 + (\nabla^2 u)_0 r^2/2n + \cdots \\ &+ (\nabla^{2p-2}u)_0 r^{2p-2}/2\cdot 4\cdots(2p-2)n(n+2)\cdots(n+2p-4), \end{aligned} \tag{12}$$

so that $A(u)$ is a polynomial in r^2.

The most general extension of these laws of the mean considered in this paper is for solutions of the differential equation

$$\sum_{i=0}^{p} c_i\nabla^{2i}u = 0 \qquad (c_p \neq 0), \tag{13}$$

where c_i are constants. In this case, the following law of the mean holds

$$\begin{vmatrix} A(u) & \phi_n(r, \lambda_1) & \cdots & \phi_n(r, \lambda_p) \\ u_0 & 1 & \cdots & 1 \\ (\nabla^2 u)_0 & \lambda_1 & \cdots & \lambda_p \\ \cdots & \cdots & \cdots & \cdots \\ (\nabla^{2p-2}u)_0 & \lambda_1^{p-1} & \cdots & \lambda_p^{p-1} \end{vmatrix} = 0, \tag{14}$$

provided that the operator on the left of (13) factors symbolically thus:

$$\sum_{i=0}^{p} c_i\nabla^{2i} = c_p \sum_{i=1}^{p} (\nabla^2 - \lambda_i), \tag{15}$$

where no two λ_i are alike. If, on the other hand, in the factorization (15) there are repeated roots, then corresponding to each root of multiplicity m, m columns of (14) past the first one are to be replaced by the λ-derivatives of the elements in (14) of order $0, 1, \cdots, m-1$. Thus, if

$$\sum_{i=0}^{p} c_i\nabla^{2i} = c_p(\nabla^2 - \lambda_1)^{m_1}(\nabla^2 - \lambda_2)^{m_2} \cdots (\nabla^2 - \lambda_k)^{m_k}, \tag{15'}$$

where $m_1+m_2+\cdots+m_k=p$ and $\lambda_i \neq \lambda_j$ for $i\neq j$, then the elements of the first row in (14) from the second one on are replaced by

$$\begin{gathered} \phi_n(r, \lambda_1), \quad \left.\frac{\partial\phi_n(r, \lambda)}{\partial\lambda}\right|_{\lambda=\lambda_1}, \cdots, \left.\frac{\partial^{m_1-1}\phi_n(r, \lambda)}{\partial\lambda^{m_1-1}}\right|_{\lambda=\lambda_1}, \\ \phi_n(r, \lambda_2), \cdots, \left.\frac{\partial^{m_k-1}\phi_n(r, \lambda)}{\partial\lambda^{m_k-1}}\right|_{\lambda=\lambda_k}, \end{gathered} \tag{16}$$

and similarly for the elements of the other rows.

An interesting application of (5) occurs in establishing the expansion

$$A(u) = \sum_{k=0}^{\infty} (\nabla^{2k}u)_0 r^{2k}/2\cdot 4 \cdots 2k\cdot n(n+2)\cdots(n+2k-2) \tag{17}$$

of the mean $A(u)$ of an arbitrary analytic function u in powers of the radius r (§4). This series will be recognized as a series whose first p terms agree with the right-hand side of (12); it can also be given the symbolic form

$$\phi_n(r, \nabla^2)u_0. \tag{18}$$

Utilizing (17), the following illuminating interpretation is derived for $\nabla^{2k}u$ at a point O:

$$(\nabla^{2k}u)_0 = \frac{2\cdot 4\cdots 2k\cdot n(n+2)\cdots(n+2k-2)}{(2k)!} A\left(\frac{\partial^{2k}u}{\partial r^{2k}}\right)_0, \tag{19}$$

where the last factor denotes the mean of the directional derivatives of order $2k$ of the function u in all directions through O, averaged over the solid angle through O (§4). A further interpretation for $\nabla^2, \nabla^4, \cdots$, also derived from (17), is given by

$$\begin{gathered} (\nabla^2u)_0 = 2n \lim_{r\to 0} [A(u) - u_0]/r^2, \\ (\nabla^4u)_0 = 2\cdot 4\cdot n(n+2) \lim_{r_2\to 0} \lim_{r_1\to 0} \frac{A_2(u) - u_0 - [A_1(u) - u_0]r_2^2/r_1^2}{r_2^4}, \end{gathered} \tag{20}$$

where $A_1(u)$, $A_2(u)$ are the means over spheres of radii r_1, r_2, and the repeated limit is obtained by letting r_1 approach zero first. The interpretations (19) and (20) are valid not merely for analytic functions u, but each one also for functions u possessing a sufficient number of continuous derivatives.

As a further application of (17) are considered in §5 the functional relations existing in certain cases between the means $A_1(u)$, $A_2(u)$ of u over "subspheres" lying in two mutually totally perpendicular flats of m, $n-m$ dimensions (§5, equations (64), (70), (71), (74), (75), (77), (79)). These include a theorem of Asgeirsson and some theorems of Bateman. The functional relations of §5 are utilized in §6 for *inverting* the averaging operation A, under certain assumptions regarding the function u.

One feature that is common to the laws of the mean (2), (4), (9), (14), as well as the indicated modification of (14), is that in every case $A(u)$ is linearly dependent upon p functions $f_1(r), \cdots, f_p(r)$, which depend on the radius r but are independent of the center O, while the coefficients of dependence, C_i, are independent of the radius r but do depend upon the position of the center O. Thus,

$$A(u) = C_1 f_1(r) + \cdots + C_p f_p(r), \tag{21}$$

or more precisely,

$$A(u)\Big|_{O,r} = C_1(O) f_1(r) + \cdots + C_p(O) f_p(r). \tag{21'}$$

This property of solutions of these differential equations actually *completely characterizes* them, as is shown by the following converse (§7):

CONVERSE THEOREM. *Let there be given a function u over a region R and of class $C^{(2p)}$ there. Let*

$$f_1(r), \cdots, f_p(r) \tag{22}$$

be p linearly independent functions of a variable r, of class $C^{(2p)}$ in r for $0 \leq r < \rho$ and whose odd derivatives $f_i^{(2k+1)}(r)$, $k = 0, \cdots, p-1$, vanish at $r = 0$. If (21') holds for a sphere of radius r and center at O for any position of the center O in R and sufficiently small radius r, then u must satisfy an equation of the form* (13) *for proper constants c_i, while the functions $f_i(r)$ must reduce to a set of p solutions of the ordinary differential equation*

$$\sum_{i=0}^{p} c_i \left(\frac{d^2}{dr^2} + \frac{n-1}{r} \frac{d}{dr} \right)^i f = 0 \tag{23}$$

* This apparent restriction on $f_i(r)$ is actually satisfied by $A(u)$ as well as by the functions $f_i(r)$ of the preceding laws of the mean.

which are analytic at $r=0$. The latter solutions are exhibited above in (16) *for the general case in which the operator on the left of* (13) *factors as in* (15′).

The differential equation (23) results from (13) if it is supposed that u depends only upon r, the distance from a fixed point. We shall refer to such solutions as "symmetric" solutions.

A particularly interesting special case of this converse is given by the following result:

If, for a function u of class $C^{(2p)}$ in R

$$A(u) = \text{polynomial of degree } p \text{ in } r^2, \tag{24}$$

for any position of the center O in R and sufficiently small radius r, then u satisfies (11) (or is p-harmonic).

The above laws of the mean and their converse involve symmetric solutions of (13) (or solutions of (23)) which are analytic at $r=0$. Symmetric solutions of (13) which are not analytic at $r=0$ occur when the function u under consideration satisfies the proper differential equation not in the complete interior of a sphere S but only over a spherical shell $R_{a,b}$ between two concentric spheres S_a, S_b of radii a, b; $0>a\geqq r\geqq b$ (§3). As an example, in a three-space, if u is harmonic in such a spherical shell $R_{a,b}$, then for the various spheres concentric with S_a, S_b the mean is given by

$$A(u) = C_1 + C_2/r, \tag{25}$$

where C_1, C_2 are constants. Even this simple result does not appear to be as familiar as its simplicity warrants.

The method of proof used in deducing the various laws of the mean is, itself, of some interest, particularly in view of its elegance and simplicity. It consists in regarding the operation which replaces u over each of a concentric family of spheres by its spherical mean $A(u)$ over that surface, as a linear functional operation A, and utilizing the permutability of the operator A with the operator ∇^2. This operator A is discussed in a preceding paper* to which we shall refer briefly as I, where its above mentioned property is proved (I, Theorem 1). We shall suppose that the reader has familiarized himself with I, at least with its introductory §1, with the definition of A and of the other operators contained in it, and with the statement of the theorems. The details of the proof of I, however, are not essential for a thorough understanding of the present paper.

Some of the above results, a search of the literature has revealed, are not

* *On operations permutable with the Laplacian*, American Journal of Mathematics, vol. 45 (1932), p. 667.

new. Thus, equations (2) and (4) have been given by H. Weber,* while a result equivalent to (17) for $n=3$ has been obtained by W. D. Niven.† These results were, nevertheless, included, because it is believed that the present proofs have decided advantages both in regard to unity and simplicity.

Many of the results of this paper can be generalized by passing from the operator A to other operators considered in I, which are likewise permutable with ∇^2. These generalizations are considered briefly in §9. A formula established in §8, due to Hobson, forms a natural transition to these generalizations. At the end of §9 are considered extensions to certain discontinuous functions.

Many interesting applications of the results of this paper can be made. Thus the integral representations of $J_0(r)$:

$$J_0(r) = \int_0^\pi e^{ir\cos\theta} d\theta/\pi = \int_{-1}^1 e^{irt}(1 - t^2)^{-1/2} dt/\pi,$$

result if one starts with the simple solution of (1) in the plane for $\lambda = -1 : u = e^{ix_1}$, and averages it over circles with center at the origin.

Again, the Laplace integral for the Legendre polynomial:

$$P_n(z) = \int_0^\pi [z + \cos\phi(z^2 - 1)^{1/2}]^n d\phi/\pi,$$

is obtained by starting with the elementary harmonic function $(x_1 + ix_2)^n$ in three-space and averaging it over circles having the x_1-axis as their axis. Similarly, the expansion

$$(1 - 2r\cos\theta + r^2)^{-1/2} = \sum r^n P_n(\cos\theta)$$

may be proved by averaging over the above circles the geometric series

$$\frac{1}{1 - (x_1 + ix_2)} = \sum (x_1 + ix_2)^n.$$

However, a systematic application of the results of this paper is reserved for a forthcoming paper entitled *On integral representation of Bessel and related functions.*

Another forthcoming paper somewhat related to the present one is entitled *Green's formulas for analytic functions.* In this paper is proved the analyticity of solutions of (13) as well as their expansibility in spherical harmonics.

---

* H. Weber, Mathematische Annalen, vol. 1 (1869), p. 7; Crelle, vol. 49 (1868), p. 222.

† W. D. Niven, Transactions of the Royal Society of London, vol. 170 (1880), p. 379.

2. **On solutions of $\sum c_i L^i(V)=0$ for general linear operators L. Application to symmetric solutions of** (13). In this section we shall obtain explicit forms for symmetric solutions of (13), that is for solutions of (23). As noted above, these solutions play an essential role in the laws of the mean considered in this paper.

Before taking up (23) and its solutions consider the equation

$$L(u) - \lambda u = 0, \tag{26}$$

where L is a linear functional operator, and λ is an arbitrary constant. It will be shown how solutions of (26) can be made to yield solutions of

$$\sum_{i=0}^{p} (c_i L^i)V = 0, \qquad c_p \neq 0, \tag{27}$$

where c_i are constants. By specializing to the case $L=d^2/dr^2+(n-1)d/dr$, (23) will be obtained. Applications to other operators L occur in §9.

The solution of (26) depends upon the parameter λ as well as upon proper independent variables; consider a solution u which is analytic in λ at $\lambda=\lambda_0$. Expanding u in powers of $\lambda-\lambda_0$:

$$u = \sum_{k=0}^{\infty} \left.\frac{\partial^k u}{\partial \lambda^k}\right|_{\lambda=\lambda_0} \frac{(\lambda-\lambda_0)^k}{k!}, \tag{28}$$

substituting in (26) written in the form

$$[(L - \lambda_0) - (\lambda - \lambda_0)]u = 0, \tag{26'}$$

applying L term-wise to the right-hand of (28), and comparing coefficients of like powers of $\lambda-\lambda_0$ on both sides, we obtain the recurrence relations (omitting the subscript in λ_0):

$$(L-\lambda)\frac{1}{k!}\frac{\partial^k u}{\partial \lambda^k} = \begin{cases} 0 \text{ for } k = 0, \\ \dfrac{1}{(k-1)!}\dfrac{\partial^{k-1}u}{\partial\lambda^{k-1}} \text{ for } k > 0, \end{cases} \tag{29}$$

provided the term-wise application of L is justifiable. From (29) it follows that

$$u, \frac{\partial u}{\partial \lambda}, \cdots, \frac{\partial^{k-1}u}{\partial \lambda^{k-1}} \tag{30}$$

are solutions of the functional equation

$$(L - \lambda)^k V = 0. \tag{31}$$

Now, consider the functional equation (27). Factoring the operator in (27), there results

$$\sum_{i=0}^{p} c_i L^i = c_p(L - \lambda_1)^{m_1} \cdots (L - \lambda_k)^{m_k}, \tag{32}$$

where $\lambda_1, \lambda_2, \cdots$ are the roots of $\sum c_i\lambda^i = 0$, and $m_1, m_2, \cdots$ their respective multiplicities. Now obviously, solutions, say, of $(L-\lambda_1)^{m_1}V=0$ are also solutions of $(L-\lambda_1)^{m_1}(L-\lambda_2)^{m_2}V=0$, and, therefore, also of (27). Hence we conclude that from the solution u of (26), the following p solutions of (27) may be obtained:

$$\left.\frac{\partial^i u}{\partial \lambda^i}\right|_{\lambda=\lambda_j}; \; i = 0, \cdots, m_j - 1; j = 1, \cdots, k. \tag{33}$$

It will be noted that the λ-differentiations in (30), (33) could be replaced by differentiation with respect to μ, where μ is a properly differentiable function of λ. Thus, the functions

$$u, \frac{\partial u}{\partial \mu}, \cdots, \frac{\partial^{k-1} u}{\partial \mu^{k-1}} \tag{30'}$$

are solutions of (31). Indeed, from the differentiation formulas

$$\frac{\partial u}{\partial \mu} = \frac{\partial u}{\partial \lambda}\frac{d\lambda}{d\mu}, \quad \frac{\partial^2 u}{\partial \mu^2} = \frac{\partial^2 u}{\partial \lambda^2}\left(\frac{d\lambda}{d\mu}\right)^2 + \left(\frac{\partial u}{\partial \lambda}\right)\frac{\partial^2 \lambda}{\partial \mu^2}, \cdots$$

it follows that the functions (30′) can be expressed linearly in terms of the functions (30).

Returning now to symmetric solutions of the differential equations (1), (8), (11), (13), we replace ∇^2 by

$$\nabla^2 = \frac{d^2}{dr^2} + \frac{n-1}{r}\frac{d}{dr} = r^{1-n}\frac{d}{dr}\left(r^{n-1}\frac{d}{dr}\right) \tag{I, (3)*}$$

thus converting them into *ordinary* differential equations. Thus (1) becomes

$$\frac{d^2u}{dr^2} + \frac{n-1}{r}\frac{du}{dr} - \lambda\mu = 0. \tag{34}$$

Two solutions of (34) are readily verified to be $\phi_n(r, \lambda)$ given by (6) and

$$\psi_n(r, \lambda) = r^{2-n}\sum_{k=0}^{\infty} (r^2\lambda)^k F_k/2\cdot 4 \cdots (2k)[(2-n)(4-n) \cdots (2-n+2k)]', \tag{35}$$

where

* As explained in §1, I, (3) refers to equation (3) of the previously cited paper I.

$$F_k = \begin{cases} 1, \text{ if } n \text{ is odd, or if } n \text{ is even and } k < (n/2) - 1, \\ \log r \text{ if } n \text{ is even and } k = (n/2) - 1, \\ \log r - \left(\dfrac{1}{2} + \dfrac{1}{4} + \cdots + \dfrac{1}{2k - n + 2}\right) \\ \qquad - \left(\dfrac{1}{n} + \dfrac{1}{n+2} + \cdots + \dfrac{1}{2k}\right) \text{ if } n \text{ is even and } k > (n/2) - 1, \end{cases}$$

and where the prime following the brackets $[(2-n)(4-n)\cdots(2-n+2k)]$ indicates that the factor zero that would occur in the indicated product for even n and $k \geqq (n/2)-1$ should be omitted. These solutions result in a natural manner when one attempts to integrate (34) by means of a power series in λ admitting term-wise r-differentiation. The solution ϕ_n is analytic for all r and λ; likewise for ψ_n except for $r=0$. We denote the coefficient of λ^k in (35) by $V_{n,k+1}(r)$ or by $V_{k+1}(r)$ if the omission of the subscript n leads to no confusion, thus:

$$\psi_n(r, \lambda) = \sum_{k=0}^{\infty} \lambda^k V_{n,k+1}(r), \tag{36}$$

$$\psi(r, \lambda) = \sum_{k=0}^{\infty} \lambda^k V_{k+1}(r). \tag{36'}$$

Applying the conclusions established above regarding solutions of the functional equation (26) to the present case, there follows for $k>1$

$$\begin{cases} \nabla^2 V_k = V_{k-1}, \\ \nabla^2[r^{2k}/2\cdot 4\cdots 2k\cdot n(n+2)\cdots(n+2k-2)] \\ \qquad\qquad = r^{2k-2}/2\cdots(2k-2)n\cdots(n+2k-4), \end{cases} \tag{37}$$

while

$$\nabla^2 V_1 = 0, \qquad \nabla^2 1 = 0; \tag{37'}$$

there also follows that

$$V_1(r), \cdots, V_p(r); \qquad 1, \cdots, r^{2p-2} \tag{38}$$

are symmetric solutions of (11). These results are obtained for $\lambda_0=0$. Similarly for general λ_0 are obtained the recurrence relations

$$(\nabla^2 - \lambda)\frac{1}{k!}\frac{\partial^k(\phi, \psi)}{\partial\lambda^k} = \frac{1}{(k-1)!}\frac{\partial^{k-1}(\phi, \psi)}{\partial\lambda^{k-1}} \text{ for } k > 0, \tag{39}$$

and the following symmetric solutions of (8):

$$\partial^i(\phi, \psi)/\partial\lambda^i; \qquad i = 0, \cdots, p - 1. \tag{40}$$

Finally, we conclude that symmetric solutions of (13) are given by (16) and by similar derivatives of $\psi_n(r, \lambda)$.

The various solutions of the proper (ordinary) differential equations just obtained may be shown to be linearly independent thus furnishing a complete set of such solutions upon which any solution would depend linearly.

For $\lambda \neq 0$ these solutions may be expressed in terms of Bessel functions. Indeed, by introducing $v = r^{n/2-1}u$ and letting $y = (-\lambda)^{1/2}r$ (where *either* determination of $(-\lambda)^{1/2}$ is used) equation (34) becomes

$$(-\lambda)r^{n/2}\left[\frac{d^2v}{dy^2} + \frac{1}{y}\frac{dv}{dy} + \left(1 - \frac{(n/2-1)^2}{y^2}\right)v\right] = 0.$$

The bracket will be recognized as the Bessel differential operator of order $m = n/2 - 1$, applied to v. Hence for $\lambda \neq 0$ solutions of (34) can be expressed linearly, say, in terms of

$$y^{-m}J_m(y), \qquad \begin{matrix} y^{-m}J_{-m}(y) \text{ for } n \text{ odd}, \\ y^{-m}\boldsymbol{Y}_m(y) \text{ for } n \text{ even}, \end{matrix} \tag{41}$$

where

$$y = (-\lambda)^{1/2}r, \qquad m = n/2 - 1.^*$$

An advantage of the solution ψ_n over the Bessel function form lies in its analyticity at $\lambda = 0$. The relation between ϕ_n and the Bessel functions is given by (7); the expression of ψ_n in terms of the latter is given by

$$\left\{\begin{aligned} \psi_n(r,\lambda) &= \Gamma(1-m)(-\lambda/2)^m\left[y^{-m}J_{-m}(y)\right] \text{ for } n \text{ odd}, \\ \psi_n(r,\lambda) &= \frac{(-\lambda)^m}{2^{m+1}\Gamma(m+1)}\left\{y^{-m}\boldsymbol{Y}_m(y)\right. \\ &\quad \left. + \left[\left(1+\frac{1}{2}+\cdots+\frac{1}{m}\right)-2-2\log\frac{(-\lambda)^{1/2}}{2}\right]y^{-m}J_m(y)\right\} \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{for } n \text{ even}, \end{aligned}\right. \tag{42}$$

where, as in (41) $y = (-\lambda)^{1/2}r$, $m = n/2 - 1$. Similarly, for non-vanishing λ, the symmetric solutions of (8) can be expressed in terms of Bessel functions, a set of solutions being furnished by

---

* E. W. Hobson, Proceedings of London Mathematical Society, vol. 25 (1893), p. 49. Hobson suggests the term "rank" for n.

The Bessel function notation is that of G. N. Watson's *Theory of Bessel Functions*, Cambridge University Press.

$$(43)\qquad y^{-m+i}\left[J_{m+i}(y),\quad \begin{matrix} J_{-m-i}(y) \text{ for } n \text{ odd} \\ Y_{m+i}(y) \text{ for } n \text{ even} \end{matrix}\right] \begin{matrix} m = n/2 - 1, \quad y = (-\lambda)^{1/2}r, \\ i = 0, \cdots, p-1. \end{matrix}$$

To prove this, replace the λ-differentiations by differentiations with respect to $\log(\lambda)^{1/2}$ (or $\log(-\lambda)^{1/2}$) thus operating on (41) with the operator $2\lambda\partial/\partial\lambda = yd/dy$, and utilize the formulas

$$(44)\qquad d[z^{-\nu}J_\nu(z)]/dz = -z^{-\nu}J_{\nu+1}(z)$$

and similar formulas for $z^\nu J_\nu(z)$ and $z^{-\nu}Y_\nu(z)$.*

It may be concluded from the above that Bessel functions of order $m=n/2-1$ and argument $y=(-\lambda)^{1/2}r$ are linearly dependent on the functions $y^m\phi_n(r,\lambda)$, $y^m\psi_n(r,\lambda)$. Hence the functions appearing in (43) and therefore also in (40) can be expressed linearly in terms of $y^{2i}\phi_{n+2i}$, $y^{2i}\psi_{n+2i}$. Thus a set of solutions of (8) for $\lambda\neq 0$ is also furnished by

$$(45)\qquad r^{2i}\phi_{n+2i}(r,\lambda),\qquad r^{2i}\psi_{n+2i}(r,\lambda);\qquad i = 0,\cdots,p-1.$$

For $\lambda\neq 0$ this set is, again, linearly independent.

The symmetric solutions of (13) can now be similarly expressed in terms of Bessel functions by means of the factorization (15′), a set of the form (43) or (45) corresponding to each repeated root.

3. **Laws of the mean for solutions of** (13). Consider a solution of (1) in a spherical shell $R_{a,b}$ or in a sphere R_b. Applying the operator A to both sides of (1) for spheres concentric with boundaries there follows

$$A(\nabla^2 u) - \lambda A(u) = 0.$$

Now, by I, Theorem 1, the first term above may be replaced by $\nabla^2[A(u)]$. Hence, $A(u)$ also satisfies (1). Since $A(u)$ is symmetric, we conclude that

$$(46)\qquad A(u) = C\phi_n(r,\lambda) + D\psi_n(r,\lambda),$$

where C, D are constants. The values of the latter depend upon the particular solution u as well as upon the position of the center.

Thus it follows from I, Theorem 1, that if u satisfies (1) in a spherical region R_b, then such is also the case with $A(u)$. Since $A(u)$ is of class C'' at $r=0$, the constant D in (46) must now vanish. Putting $r=0$ to determine the remaining constant C, we obtain

$$A(u)\Big|_{r=0} = C.$$

But $A(u)|_{r=0}=u_0$. Hence (5) results. For $n=2$ and 3 it yields (4) and (2)

* Watson, loc. cit., p. 66.

respectively. For $\lambda=0$, that is for harmonic functions, (5) reduces to the Gauss law of the mean, while (46) yields

$$A(u) = C + DV_1(r), \tag{47}$$

which reduces to (25) for $n=3$.

More generally, by applying the operator A to both sides of (13) and permuting A with $\nabla^2, \nabla^4, \cdots$ (see I, §9) one proves similarly that if u is a solution of (13) in $R_{a,b}$ or in R_b, then $A(u)$, likewise, satisfies (13) there. Being symmetric $A(u)$ must therefore reduce to a linear combination of the functions (16) and of the corresponding ψ-functions. Thus, for the general case of (13), (15)

$$A(u) = C_1\phi(r, \lambda_1) + \cdots + D_p \left.\frac{\partial^{m_k-1}\psi(r, \lambda)}{(m_k - 1)!\partial\lambda^{m_k-1}}\right|_{\lambda=\lambda_k}. \tag{48}$$

We shall determine the coefficients for the case when u satisfies (13) in R_b, that is for $r \leqq b$, $r=0$ included.

Applying $(\nabla^2-\lambda_2)^{m_2} \cdots (\nabla^2-\lambda_k)^{m_k}$ to (48) and utilizing (40) it is found that the ϕ, ψ functions corresponding to the roots $\lambda_2, \lambda_3, \cdots$ drop out. Applying products of the above operator by $(\nabla^2-\lambda_1)^{m_1-1}$, and utilizing (39), there results

$$\begin{aligned}(\nabla^2 - \lambda_1)^{m_1-1}&(\nabla^2 - \lambda_2)^{m_2} \cdots (\nabla^2 - \lambda_k)^{m_k}A(u) \\ &= (\lambda_1 - \lambda_2)^{m_2} \cdots (\lambda_1 - \lambda_k)^{m_k}[C_{m_1}\phi(r, \lambda_1) + D_{m_1}\psi(r, \lambda_1)].\end{aligned}$$

Since $A(u)$ is of class $C^{(2p)}$ at $r=0$, while the left-hand operator is of class $C^{(2p-2)}$, it follows that $D_{m_1}=0$. Similarly applying products of $(\nabla^2-\lambda_2)^{m_2} \cdots (\nabla^2-\lambda_k)^{m_k}$ by $(\nabla^2-\lambda_1)^{m_1-2}, \cdots, (\nabla^2-\lambda_1)$, one proves that $D_{m_1-1}, \cdots, D_1$ all vanish. Putting $r=0$ in the m_1 equations obtained yields successively the values $C_{m_1}, \cdots, C_1$.

A similar procedure eliminates the D's and evaluates the C's corresponding to the repeated factors $\lambda_2, \lambda_3, \cdots$. There results

$$\begin{aligned}A(u) = S\,&\frac{(\nabla^2 - \lambda_2)^{m_2} \cdots (\nabla^2 - \lambda_k)^{m_k}}{(\lambda_1 - \lambda_2)^{m_2} \cdots (\lambda_1 - \lambda_k)^{m_k}} \\ &\cdot \left\{\phi(r, \lambda_1) + \cdots + \frac{(\nabla^2 - \lambda_1)^{m_1-1}}{(m_1 - 1)!} \left.\frac{\partial^{m_1-1}\phi(r, \lambda)}{\partial\lambda^{m-1}}\right|_{\lambda=\lambda_1}\right\} u \Bigg|_{r=0},\end{aligned} \tag{49}$$

where the operators $\nabla^2, \nabla^4, \cdots$ all operate only on u (but not on ϕ, $\partial\phi/\partial\lambda, \cdots$), and S indicates a summation extended to similarly constituted terms corresponding to the repeated roots $\lambda_2, \cdots, \lambda_k$; the operators A have been suppressed from $A(u), \nabla^2A(u), \cdots$ at $r=0$ by applying $\nabla^2, \nabla^4, \cdots$ first and replacing the mean at a point by the function itself.

For special cases both the above proof and the results simplify. Thus if the roots are all alike, that is, in case (8), (49) reduces to (9). If the roots are all distinct, (49) becomes

$$\begin{aligned} A(u) = {} & \frac{(\nabla^2 - \lambda_2) \cdots (\nabla^2 - \lambda_p)}{(\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_p)} u \bigg|_{r=0} \phi(r, \lambda_1) + \cdots \\ & + \frac{(\nabla^2 - \lambda_1) \cdots (\nabla^2 - \lambda_{p-1})}{(\lambda_p - \lambda_1) \cdots (\lambda_p - \lambda_{p-1})} u \bigg|_{r=0} \phi(r, \lambda_p). \end{aligned} \tag{50}$$

To obtain the forms of these laws of the mean given by (14) and its modifications indicated in §1, we proceed directly from (48) after the ψ's have been eliminated, applying $\nabla^2, \cdots, \nabla^{2p-2}$ to both sides, putting $r=0$, utilizing

$$\nabla^{2i} \frac{\partial^j \phi(r, \lambda)}{\partial \lambda^j} \bigg|_{r=0} = \frac{\partial^j (\lambda^i)}{\partial \lambda^j}, \tag{51}$$

and eliminating C_i. The relation (51) is proved by interchanging the order of the λ- and the ∇^2-differentiations, replacing $\nabla^{2i}\phi$ by $\lambda^i\phi$, and putting $r=0$ before carrying out the λ-differentiations.

4. The spherical mean of analytic and regular functions. Interpretations of the iterated Laplacians. We shall indicate two proofs of (17) for the mean of analytic functions. Consider the function

$$u = \exp\,[c_1 x_1 + \cdots + c_n x_n] u_0$$

where $c_1, \cdots, c_n$ and u_0 are constants. This function is a solution of (1) with $\lambda = c_1^2 + \cdots + c_n^2$. Applying the law of the mean (5) to u with the center O of the spheres at the origin, we obtain for all r

$$A(\exp\,[c_1 x_1 + \cdots + c_n x_n] u_0) = \phi_n(r, \lambda) u_0, \qquad \lambda = c_1^2 + \cdots + c_n^2. \tag{52}$$

Now consider an arbitrary analytic function u; placing the origin at the center O, we may write the Taylor series formally thus:

$$u = \exp\,[x_1(\partial/\partial x_1) + \cdots + x_n(\partial/\partial x_n)] u_0, \tag{53}$$

where the exponential is to be expanded as an n-fold power series, each term multiplied by u_0, and the formal product then replaced by the corresponding partial derivative of u at the origin. Since the convergence for r less than a proper $\rho > 0$ is uniform, we may average term by term; the result is, therefore, the same as the average for the above example written as an n-fold power series in c_i where the latter are replaced by the fictitious quantities $\partial/\partial x_i$ and the resulting terms interpreted as above. Hence, for $r < \rho$ (18) follows, or, more explicitly, (17).

A more direct proof of (17) is obtained by expanding $A(u)$ in powers of r^2;

$$A(u) = C_0 + C_2 r^2 + C_4 r^4 + \cdots , \tag{54}$$

and determining the constants C_{2i} by applying ∇^{2i} to both sides, replacing $\nabla^{2i}A(u)$ by $A(\nabla^{2i}u)$, putting $r=0$, and utilizing the latter part of (37). To justify the expansion (54), write the Taylor series of u in the form

$$u = \sum_{k=0}^{\infty} \frac{1}{k!}\left(x_1 \frac{\partial}{\partial x_1} + \cdots + x_n \frac{\partial}{\partial x_n}\right)^k u_0 = \sum_{k=0}^{\infty} \frac{r^k}{k!} \left.\frac{\partial^k u}{\partial r^k}\right|_0 , \tag{55}$$

where $\partial/\partial r$ denotes differentiation along the ray from the origin through $x_1, \cdots, x_n$, with respect to the distance r from the origin. Upon taking means of both sides a series such as (54) will result, since, for odd k the contributions from opposite directions toward the mean will cancel each other.

Comparing the coefficients of the powers of r^2 in (17) with those obtained by averaging the last member of (55) we obtain (19). To derive the first relation (20), transpose the first term on the right of (17) ($k=0$) to the left, divide by r^2, and let r approach zero; the further relations (20) are derived in a similar fashion by utilizing the preceding ones.

The interpretations (19), (20), for $\nabla^2, \nabla^4, \cdots$ are proved for non-analytic functions of class $C^{(2k)}$ for proper k by replacing (53), (55) by finite sums with a remainder term of the form $o(r^{2k})$ and deriving from these a similar modification of (17).

From (17) are readily derived the formulas*

$$\int udS = K_n \sum_{k=0}^{\infty} \frac{(\nabla^{2k}u)_0 r^{2k+n-1}}{2\cdot 4 \cdots 2k\cdot n(n+2)\cdots(n+2k-2)}, \tag{56}$$

$$\int udv = K_n \sum_{k=0}^{\infty} \frac{(\nabla^{2k}u)_0 r^{2k+n}}{2\cdot 4 \cdots 2k\cdot n(n+2)\cdots(n+2k)} \tag{57}$$

for the integrals of u over the surface and the volume of S. Likewise by differentiating (17) with respect to r, then multiplying by S, one may obtain similar series for the surface integrals over S of the normal derivative of u of any order.

5. **Applications of** (17). As a further application of (17), consider the means $A_1(u)$, $A_2(u)$ of analytic functions u over "subspheres" Σ_1, Σ_2 lying in two mutually totally perpendicular flats of m, $n-m$ dimensions:

$$\Sigma_1 : x_1^2 + x_2^2 + \cdots + x_m^2 = r^2, \quad x_{m+1} = x_{m+2} = \cdots = x_n = 0, \tag{58}$$

$$\Sigma_2 : x_{m+1}^2 + \cdots + x_n^2 = r^2, \quad x_1 = x_2 = \cdots = x_m = 0. \tag{59}$$

Application of (17) (for sufficiently small r) yields

* Here, as in I, K_n denotes the "area" of a unit sphere in n-dimensions.

$$(60) \quad A_1(u) = \sum_{k=0}^{\infty} (D_1^k u)_0 r^{2k}/2\cdot 4 \cdots 2k\cdot m \cdots (m + 2k - 2),$$

$$(61) \quad A_2(u) = \sum_{k=0}^{\infty} (D_2^k u)_0 r^{2k}/2\cdot 4 \cdots 2k\cdot (n - m) \cdots (n - m + 2k - 2),$$

where

$$(62) \quad D_1 = \partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_m^2; \quad D_2 = \partial^2/\partial x_{m+1}^2 + \cdots + \partial^2/\partial x_n^2.$$

It will be shown that in certain cases there exists a linear functional relation between the means $A_1(u)$, $A_2(u)$.

Thus, if u is harmonic,

$$0 = \nabla^2 u = D_1 u + D_2 u = 0,$$

whence

$$D_1 u = - D_2 u.$$

Applying D_1 to both sides, there results

$$D_1^2 u = - D_1 D_2 u = - D_2 D_1 u = D_2^2 u,$$

and, by induction, for any k

$$(63) \quad D_1^k u = (-1)^k D_2^k u.$$

There exists thus a definite ratio between the coefficients of r^2 in (60) and (61): this proves the above statement.

Denote $A_1(u)$, $A_2(u)$ by $f_1(r), f_2(r)$. For even n and $m = n/2$ (and harmonic functions u), the functional relation reduces to

$$(64) \quad f_1(r) = f_2(ir).$$

For other cases, the relation between $f_1(r), f_2(r)$ may be obtained as follows:

Consider the series

$$(65) \quad g_p(r) = \sum_{k=0}^{\infty} \frac{r^{2k} C_k}{p(p+2)\cdots(p+2k-2)} = \Gamma\left(\frac{p}{2}\right) \sum_{k=0}^{\infty} \frac{r^{2k} C_k}{2^k \Gamma\left(\frac{p}{2} + k\right)}$$

for integer p, where the C_k do not vary with p. One may express $g_{p+q}(r)$ for $q > 0$ in terms of $g_p(r)$ as follows:

$$(66) \quad g_{p+q}(r) = \frac{\Gamma[(p+q)/2]}{\Gamma(p/2)\Gamma(q/2)} r^{2-p-q} \int_0^{r^2} s^{p-2} g_p(s)(r^2 - s^2)^{q/2-1} d(s^2).$$

This relation is even simpler if the functions

$$(67) \quad G_p(r) = r^{p-2} g_p(r)/\Gamma(p/2)$$

are introduced, whereupon (66) is replaced by

$$G_{p+q}(r) = \int_0^{r^2} G_p(s)(r^2 - s^2)^{q/2-1}d(s^2)/\Gamma(q-2). \tag{68}$$

Either (66) or (68) is readily established by termwise multiplication and integration, utilizing the Eulerian integral of the first kind; the results are valid at least in the circle $|r| < \rho$ in which g_p is analytic. The relation (68) will be recognized as an Abel integral equation; it is related to "fractional integrals"* and will be recognized as equivalent to

$$G_{p+q}(r) = \left[\frac{d}{d(r^2)}\right]^{-q/2} G_p(r). \tag{68'}$$

For even q the solution of (68) is given by

$$G_p(r) = \left[\frac{d}{d(r^2)}\right]^{q/2} G_{p+q}(r), \tag{69}$$

while for odd q the fractional integrations or differentiations cannot be eliminated, and the solution of (68) is given by

$$\begin{aligned} G_p(r) &= \left[\frac{d}{d(r^2)}\right]^{(q+1)/2}\left[\frac{d}{d(r^2)}\right]^{-1/2} G_{p+q}(r) \\ &= \left[\frac{d}{d(r^2)}\right]^{(q+1)/2}\int_0^{r^2}(r^2 - s^2)^{-1/2}G_{p+q}(s)d(s^2)/\pi^{1/2}.\dagger \end{aligned} \tag{69'}$$

To apply the above to the means $f_1(r)$, $f_2(r)$ for $m \neq n/2$, suppose that $m < n/2$, and put $p = m$, $p+q = n-m$, $g_p(r) = f_1(r)$, $g_{p+q}(r) = f_2(ir)$, (or else put $g_p(r) = f_1(ir)$, $g_{p+q}(r) = f_2(r)$). There results

$$f_2(r) = \frac{\Gamma[(n-m)/2]r^{2-n+m}}{\Gamma(m/2)\Gamma(n/2-m)}\int_0^{r^2} f_1(is)s^{m-2}(r^2 - s^2)^{n/2-m-1}d(s^2), \tag{70}$$

and, in particular, for even n and $m = n/2 - 1$,

---

* See, for instance, the author's paper on *Heaviside's operational calculus*, American Mathematical Monthly, vol. 43 (1936), pp. 332–334, 339.

† Another way of expressing the linear functional relation in question is by means of

$$g_{p'}(r) = \frac{1}{2\pi i}\int g_p(s)F\left(\frac{p'}{2}, 1, \frac{p}{2}; \frac{r^2}{s^2}\right)ds,$$

where F is the hypergeometric function, and the integration is carried out over a circle $|s| = \text{const.}$ on and within which f_p is analytic, and for $|r| < |s|$. This relation holds for any integer p, p', $p \gtrless p'$.

$$r^{n/2-1}f_2(r) = \left(\frac{n}{2} - 1\right)\int_0^r f_1(is)s^{n/2-2}ds. \tag{71}$$

The special case of (71) $n=4$ (and hence $m=1$), as well as the special case of (64), $n=4$, $m=2$, has been noted by H. Bateman.*

Suppose next, that the function u, instead of being harmonic, is an analytic solution of $D_1u=D_2u$, that is of

$$\left(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_m^2}\right)u = \left(\frac{\partial^2}{\partial x_{m+1}^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\right)u. \tag{72}$$

Equation (63) is now replaced by

$$D_1^k(u) = D_2^k(u), \tag{73}$$

and the functional relations between $f_1(r)$, $f_2(r)$ become even simpler. Thus, for even n and $m=n/2$, (64) is replaced by

$$f_1(r) = f_2(r);† \tag{74}$$

while in applying (67)–(69′), g_p, g_{p+q} are replaced by $f_1(r)$, $f_2(r)$ directly, so that for even n (69) yields

$$r^{m-2}f_1(r) = \frac{\Gamma(m/2)}{\Gamma[(n-m)/2]}\left[\frac{d}{d(r^2)}\right]^{n-2m}\left[r^{n-m-2}f_2(r)\right]. \tag{75}$$

The proof of the above results is based upon the analyticity of u and of $f_1(r)$, $f_2(r)$; this is necessarily the case with harmonic functions and their means; however, solutions of (72) need not be analytic in all the variables. That the results obtained for solutions of (72) do apply, whether they are analytic or not, follows from the fact that one may approximate to a function and its first and second derivatives by means of an analytic function and its derivatives.

Applying (74) to the one-dimensional‡ wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2} \tag{76}$$

one obtains

---

* H. Bateman, *Some geometric theorems*, etc., American Journal of Mathematics, vol. 34 (1912), pp. 332–334.

† This result is due to L. Asgeirsson, Mathematische Annalen, vol. 112 (1936).

‡ A one-dimensional "sphere" along the 1-space (x-axis) is the locus $(x-x_0)^2=r^2$, where x_0 is the center, r the radius; it consists of the two points x_0+r, x_0-r. The "spherical mean" of a function $u(x)$ over it is to be understood as $[u(x_0+r)+u(x_0-r)]/2$.

$$\frac{u(x + ct', t) + u(x - ct', t)}{2} = \frac{u(x, t + t') + u(x, t - t')}{2}. \tag{77}$$

For the three-dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2\left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2}\right) \tag{78}$$

a relation similar to (71) yields

$$A(u)\Big|_t = \int_{-r/c}^{r/c} u_0(t + t')dt'/(2r/c), \tag{79}$$

where the left-hand member is the mean of u at the time t over a sphere of radius r, while the right-hand member is the time-average of u at the center over the time interval $t-r/c$, $t+r/c$.*

6. **Inversion of the averaging process.** The relations between means over spheres in two mutually perpendicular directions considered in the preceding section are also of interest in connection with the *inverse* problem of spherical means. Of course, the operation A does not possess a unique inverse, since many different functions can give rise to the same spherical mean. However, if the function u is properly restricted, then the operation A may be inverted uniquely. Thus we may ask: what even function $f(x_1)$ of the single variable x_1, will, when averaged over spheres with center at the origin, give rise to a given function $f_2(r)$?

Suppose the f_2 is analytic in r^2 and f_1 in x^2. The following relations hold at the origin (that is at the center of the spheres):

$$\nabla^{2k} f_2\Big|_0 = \nabla^{2k}[A(f_1)]_0 = A(\nabla^{2k} f_1)_0 = (\nabla^{2k} f_1)_0 = \frac{d^{2k} f_1}{dx_1^{2k}}\Big|_0.$$

Hence between the coefficients of the expansions of f_1 in powers of x_1^2 and of f_2 in powers of r^2 exist the same ratios as between the coefficients of powers of r^2 in (60), (61) for the case (73) provided m, $n-m$ are replaced by 1 and n respectively.

Similarly, if it is supposed that n is an even function of

$$x = (x_1^2 + x_2^2 + \cdots + x_m^2)^{1/2}, \qquad m < n, \tag{80}$$

the coefficients of the expansions of $A(u)$ in powers of r^2 and of u in powers of x^2 will have the same ratios as the coefficients of the powers of r^2 in (60),

* The time $t-r/c$ corresponds to the moment at which spherical wavelets should start from the points of the sphere, so that as they diverge with the characteristic velocity c, they will arrive at the center at the time t; similarly for $t+r/c$ and converging spherical wavelets.

(61) for the case (73) provided m, $n-m$ are replaced by m, n. Carrying out the necessary modifications in (66), there results

$$A(u) = \frac{\Gamma(n/2)r^{2-n}}{\Gamma(m/2)\Gamma[(n-m)/2]}\int_0^{r^2} s^{m-2}u(s)(r^2-s^2)^{[(n-m)/2]-1}d(s^2). \tag{81}$$

The relation (81) may be established directly geometrically as follows: Break up the spheres of integration S, implied in A by means of the loci $x=$const., and let

$$y = (x_{m+1}^2 + \cdots + x_n^2)^{1/2} \tag{82}$$

and

$$ds = (dx^2 + dy^2)^{1/2} = dx\, r/y,$$

since $x^2+y^2=r^2$. The cylindrical shell between x and $x+dx$ has a base lying in $x_{m+1}= \cdots =x_n=0$ of m-dimensional content

$$dB = K_m x^{m-1}dx,$$

while upon each point of this base is "projected" an $(n-m)$-dimensional subsphere of S, of radius y and of content

$$K_{n-m}y^{n-m-1}.$$

The area of S intercepted by dB is therefore

$$dS = dB\, K_{n-m}y^{n-m-1} \sec\theta,$$

where θ is the angle between the normal to S and the projecting lines, so that $\sec\theta=r/y$. Hence

$$dS = K_m K_{n-m} x^{m-1} y^{n-m-2} r\, dx$$

and

$$A(u) = \frac{K_m K_{n-m}}{K_n r^{n-2}}\int_0^{r} u(x)x^{m-1}(r^2-x^2)^{(n-m)/2-1}dx. \tag{81'}$$

This is readily reduced to (81).

The uniqueness of the averaging process may now be based upon the known facts concerning the solution of (81).

The results of this section will be applied in the paper *On integral representations*, etc., mentioned at the end of §1.

7. **Converse theorem.** We start the proof of the converse theorem described in §1 by noting that if in

$$f_i(r) = \text{polynomial of degree } p \text{ in } r^2 + O(r^2)$$

we replace r^2 by $x_1^2 + \cdots + x_n^2$, then $f_i(r)$, regarded as a space function in a

Euclidean n-dimensional space in which r is the distance from a point P, is of class $C^{(2p)}$ even at P. Applying (19) there follows

$$\nabla^{2k} f_i(r) = b_k d^{2k} f_i(r)/dr^{2k} \Big|_{r=0},$$

where $b_k \neq 0$. Now, consider the square matrix m

$$\nabla^{2j-2} f_i(r) \Big|_{r=0}; \qquad i, j = 1, 2, \cdots, p,$$

where the Laplacians are obtained by forming space functions out of $f_1(r)$ in the manner explained; suppose, for definiteness, that i represents the order of the columns. Assume that not all the elements of the first row vanish. Then, by a reversible linear transformation on $f_i(r)$ (consisting in permuting, if necessary, two functions, then dividing the first function by its value at $r=0$, and adding a constant multiple of it to the others) it is possible to transform them into a new set of functions for which the corresponding matrix has for the elements of the first row the numbers $1, 0, \cdots, 0$. Similarly, if in the new matrix not all the elements of the second row beyond the first element vanish, it is possible to obtain p new functions related to the former ones by a reversible linear transformation and such that their matrix has for its first two rows the first two rows of the unit matrix: one needs only to transform the functions beyond the first one in a fashion similar to the above and then add to the first function a proper constant multiple of the second one. This may be continued till either

(a) the matrix m has been transformed into the unit matrix, or else,

(b) the matrix m has been transformed into a matrix whose first $p'-1$ rows, $p' \leqq p$, are rows of the unit matrix, while in row p' the principal diagonal term and all the terms following it vanish.

Denote the new functions of r into which $f_i(r)$ have thus been transformed by $F_i(r)$, and write in place of (21)

$$A(u) \Big|_{\text{center at } O} = D_1(O)F_1(r) + \cdots + D_p(O)F_p(r). \tag{83}$$

In case (a) consider (83) for an arbitrary but fixed O. Interpreting each member as a space function in the neighborhood of O, apply ∇^{2j} ($j=0, 1, \cdots, p-1$) to both sides, and put $r=0$. Replacing $\nabla^{2j} A(u)|_{r=0}$ by $A(\nabla^{2j}u)|_{r=0}$ hence by $\nabla^{2j}u|_{r=0}$, we obtain at O (hence everywhere inside R)

$$D_1(O) = u, \cdots, D_p(O) = \nabla^{2p-2}u.$$

Applying ∇^2 once more to (83), putting $r=0$, and utilizing the results obtained, we get

$$\nabla^{2p}A(u) = \nabla^{2p}u = \nabla^{2p}F_1(r)\Big|_{r=0} u + \cdots + \nabla^{2p}F_p(r)\Big|_{r=0} \nabla^{2p-2}u. \tag{84}$$

We have thus proved that u satisfies an equation of the form (13).

In case (b) we obtain in a similar way from the first p' applications of ∇^2 a differential equation for u of the form (13) but of lower "order" p'. Applying the results of §3 it follows that $A(u)$ is linearly dependent on $p'<p$ functions of r. This case therefore reduces to case (a) with a value of p lower than the value initially used.

As an example, suppose that for u of class C'', $A(u)$ is proportional to the same function $f(r)$ for any concentric spherical family irrespective of the position of the center; f is supposed to be of class C'' in r and $f'(0)=0$. Then (84) yields

$$\nabla^2 u = \left(\nabla^2 F(r)\Big|_{r=0}\right)u, \qquad F(r) = f(r)/f(0)$$

or

$$\nabla^2 u = [nf''(0)/f(0)]u.$$

In particular, if $f''(0)=0$, u is harmonic.

We close this section by considering the interesting question as to whether there exists a theorem similar to the one just proved but converse to those results of §3 for which the averages $A(u)$ are taken over spheres lying in a spherical shell $R_{a,b}$ *and enclosing the inner sphere*, so that $A(u)$ is linearly dependent on functions $f_i(r)$, not all of which are regular at $r=0$. Thus, in the plane, if u is harmonic in a ring $R_{a,b}$, the average over concentric circles of radius r and enclosing the inner boundary $r=a$ is given by $A \log r+B$, where A and B are constants whose value depends upon the position of the center. Now suppose conversely, that u is of class C'' in $R_{a,b}$ and that the average for circles of above description is given by $A \log r+B$; could one infer that u is harmonic? That such need *not* be the case can be seen from the following example.

Let $u=x_1 \log (x_1^2+x_2^2)$, so that $\nabla^2 u=4x_1/(x_1^2+x_2^2)$, $\nabla^4 u=0$; u is biharmonic but not harmonic. It will be shown that in spite of u not being harmonic, $A(u)$ is of the form $A+B \log r$ for any family of circles enclosing the origin.

Since $\nabla^4 u=0$, $A(u)=AV_1(r)+BV_2(r)+C+Dr^2$, where A, B, C, D, are constants depending upon the position of the center. Holding the latter fixed, applying ∇^2 to $A(u)$, and replacing $\nabla^2 A(u)$ by $A(\nabla^2 u)$, there results

$$A(\nabla^2 u) = BV_1(r) + 4D = B \log r + 4D.$$

Letting r become infinite, it follows that $B=D=0$, since $\nabla^2 u$ vanishes at infinity.

8. **A generalization of** (17). In this section is established the following generalization of the law of the mean (17):

$$A(uP_k) = \sum_{i=0}^{[k/2]} \frac{2^{k-2i}}{i!}\left\{\frac{\partial^{k-i}\phi_n(r, \lambda)}{\partial\lambda^{k-i}}\bigg|_{\lambda=\nabla^2}(\nabla^{2i}P_k)\left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}\right)\right\}u_0. \tag{85}$$

Here u is analytic at the origin, which point will be supposed to be the center of the spherical family implied in the averaging operation A; P_k is a homogeneous polynomial of degree k; the indicated term on the right-hand side is obtained by replacing λ by ∇^2 in the λ-power expansion of $\partial^{k-i}\phi/\partial\lambda^{k-i}$, multiplying the resulting series termwise by $\nabla^{2i}P_k$ in which $x_1, x_2, \cdots$ have been replaced by $\partial/\partial x_1, \partial/\partial x_2, \cdots$, multiplying the result by u_0, then interpreting each term by applying the indicated differentiation to u at the origin; $[x]$, as usual, denotes "the greatest integer in x."

For the case

$$P_k = H_k,$$

where H_k is harmonic, the above reduces to

$$A(uH_k) = 2^k\left\{\frac{\partial^k\phi_n(r, \lambda)}{\partial\lambda^k}\bigg|_{\lambda=\nabla^2} H_k\left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}\right)\right\}u_0, \tag{86}$$

and introducing Bessel functions by means of (10),

$$\begin{aligned} A(uH_k) = \Gamma\left(\frac{n}{2}\right) r^{k+1-n/2}\{ &(\lambda^{1/2})^{1-k-n/2}2^{(n/2-1)} \\ &\cdot I_{n/2+k-1}(r\lambda^{1/2})H_k(\partial/\partial x_1, \cdots)\}\Big|_{\lambda=\nabla^2} u_0. \end{aligned} \tag{87}$$

To prove the above, consider as in §4, the special case

$$u = \exp\,[c_1x_1 + c_2x_2 + \cdots + c_nx_n]u_0,$$

putting again

$$\lambda = c_1^2 + \cdots + c_n^2.$$

Since

$$\begin{aligned} &\nabla^2(P_k \exp\,[c_1x_1 + c_2x_2 + \cdots + c_nx_n]) \\ &\qquad = \exp\,[c_1x_1 + \cdots][\lambda^2 P_k + 2(c_1\partial/\partial x_1 + \cdots + c_n\partial/\partial x_n)P_k + \lambda P_k], \end{aligned}$$

it follows that

$$(\nabla^2 - \lambda)(P_k \exp [c_1x_1 + \cdots]) \\ = \exp [c_1x_1 + \cdots][\nabla^2 + 2(c_1\partial/\partial x_1 + \cdots)]P_k$$

and hence by induction

$$(\nabla^2 - \lambda)^i(P_k \exp [c_1x_1 + \cdots]) \\ = \exp [c_1x_1 + \cdots][\nabla^2 + 2(c_1\partial/\partial x_1 + \cdots)]^iP_k.$$

In particular

$$(\nabla^2 - \lambda)^{k+1}(P_k \exp [c_1x_1 + \cdots]) = 0.$$

Hence, the mean of $P_k\exp[c_1x_1+ \cdots]$ may be evaluated by means of (9). The last term of the right-hand sum of (9) reduces to

$$\frac{\partial^k\phi}{\partial\lambda^k}\frac{1}{k!}2^k(c_1\partial/\partial x_1 + \cdots)^kP_k\Big|_0 = \frac{\partial^k\phi}{\partial\lambda^k}2^kP_k(c_1, \cdots , c_n)$$

since P_k is a homogeneous polynomial of degree k. A similar reduction in the other terms results in

$$A(P_k \exp [c_1x_1 + \cdots]) = \sum_{i=0}^{[k/2]}\frac{2^{k-2i}}{i!}\frac{\partial^{k-i}\phi_n}{\partial\lambda^{k-i}}(\nabla^{2i}P_k)(c_1, \cdots , c_n)u_0,$$

where $[x]$ denotes the greatest integer not exceeding x, and $(\nabla^{2i}P_k)(c_1, \cdots , c_n)$ is the result of replacing the x's by the c's in the $(k-2i)$th degree polynomial $\nabla^{2i}P_k$.

Having thus established (85) for an exponential function, the case of a general analytic function is now deduced from the above, as in (24), by means of the formal representation of the Taylor series by means of an exponential.

The result (86) (for $n=3$) is due essentially to Hobson.* It will be utilized in the following section in connection with obtaining analogues of the results thus far obtained for means and the functional operator A, to other functional operators considered in I.

An interesting application of (85), (86) consists in examining the order of vanishing with r of the means $A(uH_k)$, $A(uP_k)$. It is found that

$$(88) \qquad \begin{cases} A(uH_k) = O(r^{2k}), \\ A(uP_k) = O[r^{(k,k+1)}] \text{ according as } k \text{ is (even, odd).} \end{cases}$$

More precisely,

* Hobson, Proceedings of the London Mathematical Society, vol. 24 (1892–1893), p. 80.

$$\lim_{r\to 0} \frac{A(uH_k)}{r^{2k}} = \frac{2^k}{n(n+2)\cdots(n+2k-2)} H_k\left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}\right)u_0, \tag{89}$$

$$\begin{cases} \lim_{r\to 0} \dfrac{A(uP_k)}{r^k} = \dfrac{u_0(\nabla^2)^{k/2}P_k}{2\cdot 4\cdots k\cdot n(n+2)\cdots(n+k-2)} & \text{for even } k, \\[2ex] \lim_{r\to 0} \dfrac{A(uP_k)}{r^{k+1}} = \dfrac{(\nabla^2)^{(k-1)/2}P_k\left(\dfrac{\partial}{\partial x_1}, \cdots, \dfrac{\partial}{\partial x_n}\right)u_0}{2\cdot 4\cdots k\cdot n(n+2)\cdots(n+k-1)} & \text{for odd } k. \end{cases} \tag{90}$$

9. **Extension to other operators.** Let L be any linear functional operator which is permutable with ∇^2. If u is a solution of (1), then $L(u)$ likewise satisfies (1). This follows (as in §3 for the case $L=A$) by applying L to both sides of (1) and permuting L with ∇^2:

$$L(\nabla^2 u) - \lambda L(u) = 0 = \nabla^2(L(u)) - \lambda L(u).$$

Similarly, it is shown that $L(u)$ satisfies (8), (11), or (13) if u does and if L is permutable with $\nabla^2, \nabla^4, \cdots$. In many cases this leads to results as definite as the laws of the mean proved above (properly speaking, however, the results are not covered by the title of this paper).

Consider first the operator L_k (see I, (5)):

$$L_k(u) = h_k(\omega)\int h_k'(\omega')u(r, \omega')d\omega' = h_k(\omega)I(r). \tag{91}$$

Since (see I, (7))

$$\nabla^2[h_k I(r)] = h_k\left[\frac{d^2}{dr^2} + \frac{n-1}{r}\frac{d}{dr} + \frac{k(2-n-k)}{r^2}\right]I,$$

it follows that if u satisfies (1), $I(r)$ will satisfy the ordinary differential equation obtained from (1) by replacing ∇^2 by

$$\frac{d^2}{dr^2} + \frac{n-1}{r}\frac{d}{dr} + \frac{k(2-n-k)}{r^2}; \tag{92}$$

similarly for solutions of (8), (11), (13). Thus in case of solutions of (8), $I(r)$ satisfies

$$\left[\frac{d^2}{dr^2} + \frac{n-1}{r}\frac{d}{dr} + \frac{k(2-k-n)}{r^2} - \lambda\right]^p I = 0. \tag{93}$$

Proceeding in a manner similar to that employed in §2 for the case $k=0$, the following alternative forms for solutions of (93) may be obtained

$$(94)\qquad \left.\begin{aligned} &y^{-n/2+i+1}B_{n/2+k+i-1}(y), \qquad y = (-\lambda)^{1/2}r,\\ &r^{k+2i}(\phi, \psi)_{n+2k+2i}(r, \lambda),\\ &r^k \frac{\partial^i}{\partial\lambda^i}(\phi, \psi)_{n+2k}(r, \lambda), \end{aligned}\right\} \qquad i = 0, \cdots, p-1.$$

Here $B_s(x)$ denotes a Bessel function of order s and argument x. All the three forms (91) apply for $\lambda \neq 0$; the two latter also for $\lambda = 0$. For $\lambda = 0$ the above yield the p-harmonic polynomials

$$(95)\qquad H_k,\ H_k r^2, \cdots, H_k r^{2p-2},^*$$

while from the last form (94) are obtained the p-harmonic functions

$$(96)\qquad F r^{-n-k+2+2i} h_k; \qquad i = 0, \cdots, p-1,$$

where

$$(97)\qquad F = \begin{cases} \ln r \text{ if the power of } r \text{ is even and non-negative,} \\ 1 \text{ otherwise.} \end{cases}$$

The general case (13) can now be handled by factorization (see (15), (15′)).

The operation L_k, or more explicitly, its component

$$I(r) = \int h_k'(\omega')u(r, \omega')d\omega',$$

differs only by a factor K_n/r^k from the operation $A(uH_k')$, $H_k' = h_k r^k$, considered in the preceding section.

Combining the results indicated in this section with (86) one obtains for analytic solutions of (13)

$$(98)\qquad \begin{vmatrix} \dfrac{I(r)n(n+2)\cdots(n+2k-2)}{K_n r^k} & \phi_{n+2k}(r, \lambda_1) & \cdots & \phi_{n+2k}(r, \lambda_p) \\ H_k'\left(\dfrac{\partial}{\partial x_1}, \cdots, \dfrac{\partial}{\partial x_n}\right)u_0 & 1 & \cdots & 1 \\ H_k'\left(\dfrac{\partial}{\partial x_1}, \cdots, \dfrac{\partial}{\partial x_n}\right)(\nabla^2 u)_0 & \lambda_1 & \cdots & \lambda_p \\ \cdots & \cdots & \cdots & \cdots \\ H_k'\left(\dfrac{\partial}{\partial x_1}, \cdots, \dfrac{\partial}{\partial x_n}\right)(\nabla^{2p-2} u)_0 & \lambda_1^{p-1} & \cdots & \lambda_p^{p-1} \end{vmatrix} = 0$$

---

* It is known that such products of harmonic polynomials by powers of r^2 suffice to yield a complete set of p-harmonic polynomials of any degree. See Almansi, *Sull'integrazione*, etc., Annali di Matematica, (3), vol. 2 (1899), §3.

provided that the factorization of the left-hand member of (13) has no repeated roots. For $k=0$ this reduces to (14).

Turning to other operators discussed in I, consider the operators Λ_k (see I, §7), which generalize L_k to non-integer k:

$$\Lambda_k(u) = h_k(\Omega)\int h_k'(\Omega')u(r,\Omega')d\Omega'; \tag{99}$$

here h_k, h_k' are solutions of I, (12):

$$\Delta_2 h = -k(k+n-2)h \tag{100}$$

along (Riemann surfaces spread over) the unit sphere. It is found that for solutions of (8) the differential equation (93) still applies to $I(r)$, but of course, with the proper value of k, leading to Bessel function solutions as displayed in the first form (94), but of non-integer order.

Consider next the operators $L^*_{k,n-1}$ (see I, §6):

$$L^*_{k,n-1}(u) = h_k\int_{-\infty}^{+\infty}\cdots\int h_k' u(x_1',\cdots,x_{n-1}',x_n)dx_1'\cdots dx_{n-1}' = h_k I(x_n), \tag{101}$$

where h_k, h_k' are functions of $x_1, x_2, \cdots, x_{n-1}$ satisfying the differential equation

$$\left(\frac{\partial^2}{\partial x_1^2}+\cdots+\frac{\partial^2}{\partial x_{n-1}^2}\right)h = kh. \tag{102}$$

When $L^*_{k,n-1}$ are applied under the conditions of I, Theorem IV to solutions of (13), it is found that $I(x_n)$ satisfies the equation obtained from (13) by replacing ∇^2 by

$$\partial^2/\partial x_n^2 + k. \tag{103}$$

Consider finally non-Euclidean spaces N_n and operators permutable with Δ_2, the second invariant differential operator of Beltrami. To them belong the operators L_k, L_k^*, L_k^{**} of I, §8, applied respectively over spheres, horospheres, and equidistant surfaces, the two latter in the Lobatchevsky space. Here L_k is given by (91), where r, ω are spherical coordinates, and h_k, h_k' satisfy the equation (100) over a Euclidean unit sphere. Applied to a solution of

$$\Delta_2 u - \lambda u = 0, \tag{104}$$

the operator L_k leads to a function I satisfying the equation

$$\frac{d^2I}{dr^2} + c(n-1)\cot cr\frac{dI}{dr} - \left(\frac{r^2k(k+n-2)}{\sin^2 cr}+\frac{\lambda}{c^2}\right)I = 0. \tag{105}$$

The operator L_k^* is applied over equidistant horospheres $\xi = \text{const.}$ with the element of length

$$ds^2 = e^{2c\xi}(dx_1^2 + \cdots + dx_{n-1}^2) + d\xi^2; \tag{106}$$

it is given by (101) with x_n replaced by ξ, where h, h' satisfy (100). Now, a solution of (104) leads to I satisfying

$$\frac{d^2I}{d\xi^2} + (n-1)ce^{-c\xi}\frac{dI}{d\xi} + e^{-2c\xi}kI = 0. \tag{107}$$

The operator L_k^{**} leads to (105) but with rc replaced by $\pi/2 - r/c$.

Let us map the non-Euclidean space N_n on a Euclidean sphere in E_{n+1} of radius $R = 1/c$ and define a function v in E_{n+1} thus:

$$v = R^q u. \tag{108}$$

If u satisfies (104), then

$$\begin{aligned}\nabla^2 v &= \left(\frac{\partial^2}{\partial R^2} + \frac{n}{R}\frac{\partial}{\partial R} + \frac{\Delta_2}{R^2}\right)v \\ &= [q(q+n-1) + \lambda]v/R^2.\end{aligned}$$

Hence if q is chosen as a root of

$$q(q+n-1) + \lambda = 0, \tag{109}$$

then v is harmonic in E_{n+1}. This relation will be utilized in the paper *On integral representations* etc. cited in §1.

In concluding we recall briefly the extensions of some of the above results to certain discontinuous functions. Consider, for instance, for $n=3$ the function $1/r_1$, where r_1 is the distance from a fixed point P_1. The operations A or L_k applied over spheres with center at P yield harmonic results to each side of the sphere S_1 passing through P_1. From the interpretation of $L_k(u)$ in terms of the distribution of the unit mass at P_1 over S_1 (see I, §9) follows that $I(r)$ is continuous at $r = PP_1$, but that its derivative is discontinuous there by the amount $[h_k'(\omega_p'),$ the value of h_k' at $P_1]/4\pi$. A convenient way of proving this and other similar results is by spreading out the concentrated (point) mass over a finite region, thus representing $1/r$ as the limit of a function with a continuous Laplacian.

General Electric Company,
Schenectady, N. Y.

THE SCHÖNEMANN-EISENSTEIN IRREDUCIBILITY CRITERIA IN TERMS OF PRIME IDEALS*

BY

SAUNDERS MACLANE

1. **Introduction.** The Eisenstein criterion† for the irreducibility of a polynomial has been repeatedly generalized, in many cases by the use of Newton polygons. All of these irreducibility criteria for polynomials can be systematically viewed in terms of non-archimedean absolute values—so that we can state a general theorem which includes all these theorems as special cases and which also establishes the irreducibility of new classes of polynomials. Our general theorem asserts, in effect, that a polynomial $G(x)$ with no multiple roots and with rational coefficients is irreducible if there is a rational prime p which has just one prime ideal factor in the ring $R[x]/G(x)$, which is obtained by reducing modulo $G(x)$ the ring of all polynomials with rational coefficients. This criterion can be constructively applied by using a previously developed method for actually exhibiting the prime decomposition of any p.‡

The known irreducibility criteria are simply conditions which imply that the first few stages of the prime ideal construction will show that p has but one prime ideal factor. The Schönemann criterion asserts the irreducibility of polynomials of the form

$$f(x) = \phi(x)^e + pM(x), \tag{1}$$

where $\phi(x)$ is an irreducible polynomial modulo p and where $M(x)$ is a polynomial relatively prime to ϕ, mod p, and of degree less than the degree of f. Alternatively, these conditions show that in the ring $R[x]/f(x)$, p has just one prime ideal factor $P=(p, \phi(x))$, and that this factor may be found by the "second stage" of the construction of the factors of p. Because there is but one prime factor,§ and because the degree of P times the exponent to which P divides p is the degree of $f(x)$, $f(x)$ must be irreducible.

* Presented to the Society, January 2, 1936; received by the editors February 23, 1937.

† For a simple statement see B. L. van der Waerden, *Moderne Algebra*, §22.

‡ S. MacLane, *A construction for absolute values in polynomial rings*, these Transactions, vol. 40 (1936), pp. 363–395; S. MacLane, *A construction for prime ideals as absolute values of an algebraic field*, Duke Mathematical Journal, vol. 2 (1936), pp. 492–510. We refer to these two papers as Const I and Const II, respectively. They contain the definition of absolute values, etc., used subsequently.

§ The connection between the Eisenstein irreducibility criterion and the prime ideal factorization of a rational prime was observed by M. Bauer, *Zur allgemeinen Theorie der algebraischen Grössen*, Journal für die Mathematik, vol. 132 (1907), pp. 21–32, especially §IV; also by O. Perron, *Idealtheorie und Irreduzibilität von Gleichungen*, Mathematische Annalen, vol. 60 (1905), pp. 448–458.

Our new irreducibility criterion may be stated with reference to a rational prime p or, alternatively, in terms of the corresponding "p-adic" absolute value. This simple form of the theorem is stated in §2 for a polynomial with coefficients in any field K. It involves certain absolute values of the polynomial ring $K[x]$. We include also a more general theorem giving all possible degrees for the factors of a reducible polynomial.* Next, in §3, we indicate how our result includes both old and new cases. To establish the prime ideal interpretation, we first develop briefly in §4 the properties of prime ideals in a ring $K[x]/G(x)$, where K is an algebraic number field. These properties give the irreducibility theorem in the prime ideal form. Finally in §5 we show how the successive "approximant" values used in our irreducibility criteria do, in fact, give a construction for the prime ideals in the corresponding ring $K[x]/G(x)$. Hence the general irreducibility theorem, stated in terms of absolute values, implies the form of the irreducibility theorem already stated in terms of prime ideals.

The fundamental irreducibility theorem of §2 can also be applied to polynomials in several variables. In the last section we give several specific examples of the new irreducibility criteria which result.

2. **Irreducibility criteria with approximants.** An *absolute value* of a ring is a function $V(a)$ defined for all a in the ring and with the properties

$$V(ab) = V(a) + V(b). \qquad V(a+b) \geqq \min (Va, Vb).$$

An element a of the ring is *equivalence-divisible* in V by an element b if there is an element c with $V(a-bc) > V(a) = V(bc)$.

Consider now polynomials with coefficients in any field K. A polynomial $f(x)$ is a *key polynomial* over a value V of $K[x]$ if $f(x)$ has the first coefficient 1, if any polynomial equivalence-divisible by $f(x)$ in V has a degree at least as great as the degree of $f(x)$, and if any product equivalence-divisible by $f(x)$ in V has a factor equivalence-divisible by $f(x)$ in V.† The first form of our general irreducibility criterion for polynomials is

THEOREM 1. *If K is any field and if $G(x)$ is a key polynomial over a value V of the polynomial ring $K[x]$, then $G(x)$ is irreducible.*

Proof. Suppose that $G(x)$ could be factored as $G(x) = f(x)h(x)$. Then this product is equivalence-divisible by G in V. As G is equivalence-irreducible, by the definition of a key, one of the factors f or h must be equivalence-

* Simple theorems of this type have been stated by Dumas and Ore (cf. §3 below) and by O. Perron, loc. cit., and E. Netto, *Ueber die Irreductibilität ganzzahliger ganzer Funktionen*, Mathematische Annalen, vol. 48 (1896), pp. 82–88.

† MacLane, Const I, Definition 4.1.

divisible by G. But G is also minimal, so that this factor has a degree at least that of G. Therefore the assumed decomposition is trivial, so G is irreducible.

The relevance of this Theorem derives from the possibility of explicitly constructing all possible values of $K[x]$ for many fields K and from an explicit criterion* which determines when $G(x)$ is a key polynomial over such values. Any value V in $K[x]$ determines a value $V_0a = Va$ in the coefficient field K. The simplest values of $K[x]$ are the "inductive values"† V_k characterized by the properties: (i) V_k agrees in the field K with a given value V_0; (ii) V_k gives certain polynomials $\phi_1(x) = x, \phi_2(x), \cdots, \phi_k(x)$ specific assigned values $V_k\phi_i(x) = \mu_i$; (iii) V_k assigns to every polynomial in the ring $K[x]$ the smallest possible value consistent with the conditions (i) and (ii). Such an inductive value is denoted by

$$V_k = [V_0, V_1x = \mu_1, V_2\phi_2(x) = \mu_2, \cdots, V_k\phi_k(x) = \mu_k]. \tag{2}$$

These values may be obtained by an inductive definition of the values V_i of $K[x]$ determined by the first i polynomials $\phi_i(x)$. In (2), each μ_i is a number, while each polynomial $\phi_i(x)$, with $i > 1$, must be a key polynomial over the previous inductive value V_{i-1}, and must satisfy two other minor conditions (Const I, Definition 6.1). The set of all numbers $\nu = V_kf(x) - V_kg(x)$ which are values of rational functions $f(x)/g(x)$ in V_k is a group, the *value-group* Γ_k. For a given $\phi_k(x)$ there is for each polynomial $G(x)$ a unique "expansion" in powers of ϕ_k, of the form

$$G(x) = g_m(x)\phi_k^m + g_{m-1}(x)\phi_k^{m-1} + \cdots + g_0(x), \tag{3}$$

where the coefficients $g_i(x)$ in the expansion are 0 or of degree less than the degree of ϕ_k. The value V_kG is the minimum of the values of the terms in this expansion.

The hypothesis in Theorem 1 that G is a key polynomial is most easily fulfilled by making a suitable multiple of G have a residue class modulo V (Const I, part II) which is a linear polynomial.

THEOREM 2. *If V_k is an inductive value with a last key polynomial ϕ_k, and if a polynomial $G(x)$ has an expansion* (3) *in terms of ϕ_k such that* (i) $g_m(x) = 1$; (ii) $V_kG = V_k\phi_k^m = V_kg_0$; (iii) *if $n < m$ is a positive integer, then $n\mu_k$ is not in the value-group Γ_{k-1}; then $G(x)$ is a key polynomial over V_k and hence irreducible.*

---

* The method of Const I, Theorem 9.4 and §13 applies whenever V is an inductive value and K an algebraic field.

† For the explicit definition, see MacLane, Const I, §4, (3) or Const II, §2, (4).

Proof. Conditions (i) and (ii) will make $G(x)$ a key, by Const I, Theorem 9.4, provided we also show $G(x)$ equivalence-irreducible in V_k. But any G has by Const II, Theorem 4.2 a representation as a product of key polynomials $\psi(x)$. Each one must have the form of an expansion

$$\psi(x) = \phi_k^e + h_{e-1}(x)\phi_k^{e-1} + \cdots + h_0(x), \qquad e > 0$$

in which the first and last terms again have the same value in V_k. By the minimal property (iii) of m this is possible only if e is a multiple of m. Hence the representation of G has just one factor ψ. This factor has the same degree and the same equivalence-divisors as G, so that G, like ψ, is equivalence-irreducible and therefore a key polynomial.* These theorems include and generalize all the classical irreducibility criteria of the Newton polygon type.

To obtain information about the degree of possible factors of a reducible polynomial, we use "approximants." If V_k is a finite and homogeneous† kth stage value of $K[x]$, and if $G(x)$ is any polynomial expanded as in (3), consider the exponents j in (3) with the property that $V_kG = V_k(g_j(x)\phi_k^j)$, and denote by α the largest and by β the smallest of these exponents. Then the *projection* of G on V_k is taken to be

$$\text{proj}\,(V_k, G) = \alpha - \beta. \tag{4}$$

The homogeneous value V_k, considered as an extension of the value V_0 of K, is an *approximant* of G over V_0 if and only if proj $(V_k, G) > 0$. For any polynomial $G(x)$ with coefficients in an algebraic number field, the set of all kth approximants can be found by Newton polygons with the procedure used in Const II for an irreducible polynomial $G(x)$.

A homogeneous inductive value V_s is non-finite if $V_s\phi_s = \mu_s = \infty$. We call such a V_s an *improper kth approximant* to $G(x)$, for any $k \geqq s$, if ϕ_s is a factor of $G(x)$. We define proj (V_s, G) as the exponent to which ϕ_s divides G. Henceforth the phrase "kth approximants" refers both to proper and improper kth approximants.

To interpret (4), note that the integer β can be uniquely characterized by the properties

$$\text{(4a)} \qquad \begin{array}{l} \phi_k^\beta \text{ is an equivalence-divisor of } G(x) \text{ in } V_k, \\ \phi_k^{\beta+1} \text{ is not an equivalence-divisor of } G(x) \text{ in } V_k. \end{array}$$

---

* This could also have been proven by the method of Const I, by finding an $R(x)$ so that $R \cdot G$ has value 0, showing (Theorem 12.1) that $R \cdot G$ has a residue-class which is a linear polynomial and thus proving G equivalence-irreducible by Lemma 11.2.

† Homogeneity is required because every inductive value with a discrete V_0 is equal to one and only one homogeneous value (cf. Const I, p. 393).

For all terms to the right of $g_\beta(x)\phi_k^\beta$ in the expansion (3) certainly have larger values than this term, hence the first half of (4a). If in addition the second half of (4a) were false, there would be an $h(x)$ with $f(x)\sim h(x)\phi_k^{\beta+1}$ in V_k; that is, with

$$V_k[f(x) - h(x)\phi_k^{\beta+1}] > V_kf(x) = V_k[h(x)\phi_k^{\beta+1}].$$

By inserting in $f=h\phi_k^{\beta+1}+[f-h\phi_k^{\beta+1}]$ the expansions in powers of ϕ_k for $h(x)$ and for the term in brackets we get an expansion for $f(x)$ in which the last term with the minimal value has the exponent $\beta+1$ or greater, counter to the definition of β. Therefore (4a) characterizes* β.

For a product $G=f'(x)f''(x)$, each factor f', or f'', has exponents α' and β', or α'' and β'', as in (4). By (4a), β' is the power to which ϕ_k is an equivalence-divisor of f', and similarly for β''. Therefore, by the uniqueness of the equivalence decomposition, $\beta=\beta'+\beta''$ (see Const II, §4). Furthermore α is the "effective degree" of $G(x)$ and has the property $\alpha=\alpha'+\alpha''$ (see Const II, §4, (1)). Combining these, we have

$$\text{proj}\ (V_k, f'(x)f''(x)) = \text{proj}\ (V_k, f'(x)) + \text{proj}\ (V_k, f''(x)). \tag{5}$$

This also holds for improper approximants. We obtain at once

LEMMA 1. *The kth approximants of a product $G(x)H(x)$ over V_0 consist of all the kth approximants of $G(x)$ and of all the kth approximants of $H(x)$ over V_0.*

THEOREM 3. The degree-of-factors theorem. *If for a value V_0 of the field K all the kth stage homogeneous approximants to the polynomial $G(x)$ over V_0 are denoted by $V_k^{(i)}$, $i=1, \cdots, t_k$, then any factor $f(x)$ of $G(x)$ has a degree*†

$$\deg f(x) = \sum_{i=1}^{t_k} c_i \deg \phi(V_k^{(i)}) \tag{6}$$

where each c_i is any number satisfying

$$0 \leqq c_i \leqq \text{proj}\ (V_k^{(i)}, G), \qquad (i = 1, 2, \cdots, t_k).$$

Proof. This will follow at once by (5) combined with the equation

$$\deg G(x) = \sum_{i=1}^{t_k} \text{proj}\ (V_k^{(i)}, G)\cdot \deg \phi(V_k^{(i)}) \tag{7}$$

for any polynomial $G(x)$. To establish (7), decompose $G(x)$ into its irreducible

---

* A longer proof of this fact is also given in Const II, Theorem 5.1.

† Here $\deg \phi(V_k^{(i)})$ represents the degree of the last key polynomial of $V_k^{(i)}$. It is equal to the degree of the residue-class field of V_k multiplied by the exponent of V_{k-1}. See Const II, §9, (1).

factors, apply Const II, Theorems 5.2 and 5.3* to obtain equations similar to (7) for each such irreducible factor, and combine these equations by using the relation (5).

If $G(x)$ has no multiple factors and V_0 is "discrete," then k can be taken so large that every kth approximant has† the projection 1. A consequence is:

THEOREM 4. *If for a value V_0 of K and an integer k there is only one kth stage approximant V_k to $G(x)$, and if* proj $(V_k, G)=1$, *then $G(x)$ is irreducible.*

If V_k be replaced by V_{k+1}, this is essentially a restatement of Theorem 1 with the additional advantage that the value V introduced arbitrarily there is now characterized as an approximant of $G(x)$.

Irreducibility can also be established by several applications of the degree-of-factors theorem.

THEOREM 5. *If V_0 and W_0 are two given values of a field K, if $G(x)$ is of degree $n=n_1\cdot n_2$, where $(n_1, n_2)=1$, and if there is an integer k such that every kth approximant V_k to $G(x)$ over V_0 has* deg $\phi(V_k)\equiv 0 \pmod{n_2}$, *while every kth approximant W_k to $G(x)$ over W_0 has* deg $\phi(W_k)\equiv 0 \pmod{n_1}$, *then $G(x)$ is irreducible.*

Proof. By the condition on V_0 and Theorem 3 each factor $f(x)$ has a degree which is a sum of multiples of deg $\phi(V_k)$ and which is therefore a multiple of n_2. By the same argument for W_0 the degree of $f(x)$ is a multiple of n_1, and so this degree must be n, the degree of $G(x)$.

This theorem can be generalized to apply to s different values with $n=n_1n_2\cdots n_s$. The conditions on the approximants over V_0 can be fulfilled, for example, by making the first approximant have the exponent n_2, for then deg $\phi(V_2)$ must be a multiple of n_2.

3. **Examples.** The theorem of Schönemann, as stated in §1, (1), follows from our results, for the condition on $\phi(x)$ is sufficient to make ϕ a key polynomial for a value

$$V_2 = [V_0 p = 1, V_1 x = 0, V_2\phi(x) = 1/e],$$

while the condition on $M(x)$ makes the last term in the expansion of $f(x)$ in powers of ϕ have the value 1. Hence $f(x)$ is irreducible by Theorem 2.

In a similar fashion, the various generalizations of the Eisenstein irreducibility theorem for polynomials with rational coefficients can all be shown

---

* It may be observed that the assumption that V_0 was discrete was made in §5 only to insure that the approximants and limit values give all possible values of $K[x]$, so that this discreteness assumption is not needed here.

† By Const II, Theorem 8.1, which applies whenever G has a non-vanishing discriminant. The b_i in this theorem are then all 0 or 1.

to depend on the use of absolute values which are the extension of p-adic values. This is illustrated in the following list of known theorems which are special cases of our theorem applied to particular first and second stage values. Unless otherwise indicated the author stated our Theorem 2 for the field of rational numbers and for the special value indicated, in most cases not explicitly in terms of absolute values but in some equivalent form.*

Eisenstein:	$[V_0p=m,\ V_1x=1]$.
Schönemann:	$[V_0p=m,\ V_1x=0,\ V_2\phi=1]$.
Königsberger:	$[V_0p=n,\ V_1x=r]$, $r>0$. Also Theorem 5 for two such first stage values.
Bauer:	$[V_0p=n,\ V_1x=0,\ V_2\phi=\alpha]$.
Dumas:	$[V_0p=n,\ V_1x=r]$, Newton polygons, Theorem 3.
	$[V_0p=n,\ V_1x=r,\ V_2g=s]$, with restrictions.
Kürschák:	$[V_0p=1,\ V_1x=\mu,\ V_2f=\nu]$, Newton polygons, $V_1f=0$.
Rella:	As in Kürschák, for K a domain of integrity.
Ore:	$[V_0p=n,\ V_1x=0,\ V_2\phi=r]$, Newton polygons, Theorem 3.
Ore:	$[V_0p=n,\ V_1x=r]$, Theorem† 1.
Ore:‡	$[V_0p=n,\ V_1x=0,\ V_2\phi_2=r,\ V_3\phi_3]$. The degree-of-factors Theorem 3.

Irreducibility criteria have also been systematized by Blumberg§ in terms of a notion of "rank." This "rank" is closely related to our "absolute value." It applies also to differential expressions, but does not include the higher stage values.

Our methods for constructing inductive values allow the construction of

* The papers cited here are, in order: G. Eisenstein, *Ueber die Irreduzibilität und einige andere Eigenschaften der Gleichungen*, etc., Journal für die Mathematik, vol. 39 (1850), p. 166; Th. Schönemann, *Von denjenigen Moduln, welche Potenzen von Primzahlen sind*, Journal für die Mathematik, vol. 32 (1846), pp. 93–105, §61; L. Königsberger, *Ueber den Eisensteinschen Satz von der Irreduzibilität algebraischer Gleichungen*, Journal für die Mathematik, vol. 115 (1895), pp. 53–78, especially (67) on p. 69; M. Bauer, *Verallgemeinerung eines Satzes von Schönemann*, Journal für die Mathematik, vol. 128 (1905), pp. 87–89; G. Dumas, *Sur quelques cas d'irreductibilité des polynomes à coefficients rationnels*, Journal de Mathematique, (6), vol. 2 (1906), pp. 191–258; J. Kürschák, *Irreduzible Formen*, Journal für die Mathematik, vol. 152 (1923), pp. 180–191; T. Rella, *Ordnungsbestimmungen in Integritätsbereichen und Newtonsche Polygone*, Journal für die Mathematik, vol. 158 (1927), pp. 33–48; O. Ore, *Zur Theorie der Irreduzibilitätskriterien*, Mathematische Zeitschrift, vol. 18 (1923), pp. 278–288; O. Ore, *Zur Theorie der Eisensteinschen Gleichungen*, Mathematische Zeitschrift, vol. 20 (1924), pp. 267–279.

† This is the first treatment of a non-linear criterion.

‡ O. Ore, *Zur Theorie der algebraischen Körper*, Acta Mathematica, vol. 44 (1924), pp. 219–314. Theorem 4 on page 230 is stated for an algebraic number field as coefficient field, while Theorem 9 on page 240 gives the degree of any factor in terms of the first two stages, plus the key polynomials only on the third stage. Hence this theorem differs in form from our statement.

§ H. Blumberg, *On the factorization of expressions of various types*, these Transactions, vol. 17 (1916), pp. 517–544.

examples of polynomials which are irreducible in virtue of arbitrary complicated inductive values not falling under the above cases. For example, the value

$$[V_0p = 4, V_1x = 0, V_2(x^2+1) = 2, V_3((x^2+1)^2+p) = 5] \qquad (p = 3)$$

of the ring of polynomials with rational coefficients proves

$$f(x) = [(x^2+1)^2+p]^2 + p^2(x^2+1) = x^8 + 4x^6 + 12x^4 + 25x^2 + 25$$

irreducible by Theorem 2, although the second stage approximant V_2 does not show it irreducible. The use of non-homogeneous key polynomials (or of constant degree inductive values) is illustrated by

$$[V_0p = 2, V_1x = 0, V_2(x^2+1) = 2, V_3(x^2+1+p) = 3] \qquad (p = 7),$$

which, by Theorem 2, proves the irreducibility of

$$(x^2+1+p)^2 + p^3 = x^4 + 16x^2 + 407.$$

A case of Theorem 1 not included in the linear Theorem 2 is

$$V_2 = [V_0p = 1, V_1x = 0, V_2(x^2+x+1) = 1] \qquad (p = 5),$$

$$\begin{aligned} f(x) &= (x^2+x+1)^2 - p(x^2+x+1) + 3p^2x + 3p^3 \\ &= x^4 + 2x^3 - 2x^2 + 72x + 371. \end{aligned}$$

The residue-class ring of $K[x]$ for this V_2 is by Const I, Theorem 12.1 just the ring $F[\theta, y]$, where F is the field of integers, modulo 5, and θ is the residue-class of x and so is the root of x^2+x+1 over F, while y is a symbol representing the residue-class* of $(x^2+x+1)/p$. To test this $f(x)$ for equivalence-irreducibility we first multiply it by p^{-2} to make it have the value 0. Then

$$p^{-2}f(x) = [(x^2+x+1)/p]^2 - [(x^2+x+1)/p] + 3x + 3p,$$

so that the residue-class polynomial is

$$H_2[p^{-2}f] = y^2 - y + 3\theta$$

which, although not linear, is irreducible over the Galois field $F(\theta)$. Hence $F(x)$ is equivalence-irreducible† and therefore irreducible by Theorem 1.

4. **Prime decomposition in algebraic rings.** To interpret our irreducibility criteria for $G(x)$ we first summarize some arithmetic properties of the corresponding residue-class ring

$$A = K[x]/(G(x)). \tag{8}$$

* See Const I, §12, (6).

† Const I, Lemma 11.2, where $R = p^{-2}$.

Here, and throughout §§4 and 5, K denotes an algebraic number field, $G(x)$ is a polynomial in $K[x]$, and $\mathfrak{O}$ the ring of all integers of the commutative algebra A.

THEOREM 6. *In $\mathfrak{O}$, every ideal B which is not a divisor of zero* has a decomposition, unique except for the order of factors, as a product of prime ideals from $\mathfrak{O}$. For ideals B and C, B not a divisor of zero, the inclusion $B \subset C$ implies the existence of an ideal D with† $B = CD$. Let*

$$G(x) = g_1(x)^{e_1} g_2(x)^{e_2} \cdots g_t(x)^{e_t} \tag{9}$$

be the decomposition of $G(x)$ into distinct irreducible factors $g_i(x)$, and denote by K_i the algebraic field $K[x]/(g_i(x))$, and by $\mathfrak{O}_i$ the ring of all the integers of K_i. For each prime ideal $P_i \neq \mathfrak{O}_i$ of the ring $\mathfrak{O}_i$ there is a corresponding prime ideal P' of $\mathfrak{O}$, and the residue-class rings $\mathfrak{O}/P'$ and $\mathfrak{O}_i/P_i$ are isomorphic. These ideals P' are all distinct and include all the prime ideals of $\mathfrak{O}$, except for $\mathfrak{O}$ itself. If $\mathfrak{p}$ is a prime ideal of the ring of integers of the base field K, the decomposition of $\mathfrak{p}$ in $\mathfrak{O}$ may be found by decomposing $\mathfrak{p}$ in each $\mathfrak{O}_i$, replacing each prime factor P_i in these decompositions by the corresponding P' and multiplying the resulting decompositions.

The proof is omitted, since the results are implicit in the more general arithmetic of non-commutative algebras.‡ The theorem can easily be obtained directly by the usual consideration of A as the direct sum of the fields K_i.

The degree of a prime ideal P' in the ring of integers $\mathfrak{O}$ of Theorem 6 is defined to be the degree of its residue-class ring $\mathfrak{O}/P'$ over the residue-class ring of $\mathfrak{p}$, where $\mathfrak{p}$ is the prime ideal of K such that $\mathfrak{p} \cdot \mathfrak{O} \subset P'$. This is, by the theorem, the same as the degree of the corresponding prime ideal P_i of K_i over $\mathfrak{p}$. The relation between the degree of an algebraic number field and the degree and exponents of prime ideals yields then the following analogue to our irreducibility theorem.

THEOREM 7. The degree-of-factors theorem. *If $G(x)$ has no multiple factors and if, in the ring of integers of the algebraic number field K, $\mathfrak{p}$ is any prime ideal with the decomposition*

* An ideal B in $\mathfrak{O}$ is a divisor of zero if every element of B is a divisor of zero.

† This second property, "every divisor is a factor," is closely associated with the decomposition into prime ideals (van der Waerden, *Moderne Algebra*, vol. 2, §100). When it holds for all ideals, including divisors of zero, the ring is called a multiplication-ring. See Krull, *Idealtheorie*, Ergebnisse der Mathematik, vol. 5, p. 26.

‡ M. Deuring, *Algebren*, Ergebnisse der Mathematik, vol. 4, p. 108. E. Artin, *Zur Arithmetik hyperkomplexer Zahlen*, Abhandlungen des Mathematischen Seminars, Hamburg, vol. 5 (1928), pp. 261–289.

$$\mathfrak{O}\cdot\mathfrak{p} = P_1^{b_1}P_2^{b_2}\cdots P_s^{b_s} \qquad (b_i = \exp P_i) \tag{10}$$

into prime ideals in $\mathfrak{O}$, *then any factor* $f(x)$ *of* $G(x)$ *has a degree of the form*

$$\deg f(x) = \sum{}' b_i\cdot\deg P_i = \sum{}'(\exp P_i)(\deg P_i), \tag{11}$$

where the sum is to be taken over any subset of the given set of prime ideal factors P_i.

THEOREM 8. The irreducibility criterion. *If* $\mathfrak{p}$ *is a prime ideal from the algebraic number field* K *and if* $\mathfrak{p}$ *has but one prime ideal factor in* $\mathfrak{O}$, *then the polynomial* $G(x)$ *is a power of an irreducible polynomial.*

Proof. The decomposition of $\mathfrak{p}$ in $\mathfrak{O}$ is obtained, as in Theorem 6, by combining decompositions from all the direct summand fields K_i. If the final decomposition of $\mathfrak{p}$ is to have but one factor, there can be only one such direct summand and hence only one irreducible factor $g_i(x)$ in (9).

THEOREM 9. *If* K *is an algebraic number field,* δ *an integer in* K, *and* $G(x)$ *a polynomial such that the principal ideal* (δ) *becomes the nth power of an ideal in* $\mathfrak{O}$, *then any factor of* $G(x)$ *in* $K[x]$ *has a degree* r *such that* $(\delta)^r$ *is the nth power of some ideal in the ring of integers of* K.

One case of this theorem, in which the decomposition $(\delta)=B^n$ was insured by specifying the form of $G(x)$ and taking $n=\deg G(x)$, was first stated by Sopman.* The general Theorem 9 can be established by Sopman's methods, applied to the fields K_i, or by the direct use of Theorem 8.

5. **Approximants and prime ideals in algebraic rings.** Our two forms of the irreducibility criterion are essentially the same, because the approximants to any polynomial $G(x)$ can be used to construct the prime ideals in the corresponding $K[x]/(G(x))=A$. For a prime ideal $\mathfrak{p}$ from the algebraic number field K let V_0 be the $\mathfrak{p}$-adic absolute value of K. Without essential loss of generality, we assume throughout this section that $G(x)$ has first coefficient 1 and its other coefficients V_0-integers. If in (9) we make all factors $g_i(x)$ have the first coefficient 1, they will also have V_0-integers as coefficients.

THEOREM 10. *Given a* $G(x)$ *and a* $\mathfrak{p}$*-adic value* V_0 *of* K, *there is for* k *sufficiently large a one-to-one correspondence between the kth approximants* V_k *to* G *over* V_0, *and the prime ideal factors* P' *of* $\mathfrak{p}$ *in* $\mathfrak{O}$. *If* ϕ_k *is the last key polynomial of* V_k, *and* P' *is the corresponding prime ideal, then*

$$\deg \phi_k = (\deg P')(\exp P'). \tag{12}$$

* M. Sopman, *Ein Kriterium für Irreduzibilität ganzer Funktionen in einem beliebigen algebraischen Körper*, Mathematische Annalen, vol. 91 (1924), pp. 60–61.

Proof. If for G as in (9) we set $G^*(x)=g_1(x)g_2(x)\cdots g_t(x)$, then, by Lemma 1, G and G^* have the same kth approximants. The "finiteness" Theorem 8.1 of Const II applies to G^*, and gives a k' so large that, for $k\geqq k'$, every kth approximant V_k to G^* has the projection 1. Each V_k is then, by (5), an approximant to just one factor $g_i(x)$ of G^*. But each approximant to the irreducible $g_i(x)$ constructs an extension W of the value V_0 to the field K_i (Const II, Theorem 10.2). Each such W corresponds to just one prime ideal factor P_i of $\mathfrak{p}$ in $\mathfrak{O}_i$, while, by Theorem 6, P_i corresponds to just one prime ideal factor P' of $\mathfrak{p}$ in $\mathfrak{O}$. Combining these successive correspondences we find that approximants V_k do correspond to prime ideals P', as asserted. The relation (12) follows by Const II, Theorem 9.3.

With this connection between the approximants and prime ideals, the two forms of the two irreducibility criteria become essentially identical. For if $G(x)$ has no multiple roots, then in the presence of Theorem 10, the irreducibility criterion of Theorem 8 in terms of prime ideals is immediately equivalent to the irreducibility criterion of Theorem 4 in terms of approximants. Similarly results hold for the degree-of-factors theorem, if k is large, for then (12) reduces (6) to (11). Hence the generalizations of the Eisenstein criterion are merely statements about prime ideal decompositions.

6. **Irreducibility of polynomials in several variables.** A number of theorems concerning such polynomials with coefficients in any field F will now be derived from the general results of §2.

THEOREM 11. *If $\phi(x)$ is irreducible over the field F, if $g(x, y)$ is a polynomial in $F[x, y]$ of degree in y less then n, and if $a(x)\not\equiv 0 \pmod{\phi}$ and $g(x, 0)\not\equiv 0 \pmod{\phi}$, then*

$$f(x, y) = a(x)y^n + g(x, y)\phi(x)$$

is irreducible in $F[x, y]$, except perhaps for a factor involving only x.

Proof. By the irreducibility of $\phi(x)$

$$V_0 = [V_0F = 0,\ V_0x = 0,\ V_0\phi(x) = 1]$$

is a value of the coefficient field $F(x)$, where we have indicated by $V_0F=0$ that $V_0a=0$ for all constants $a\neq 0$ in F. The inductive values $V_1=[V_0, V_1y=1/n]$ is an approximant to $f(x, y)$, and the irreducibility then follows by Theorem 2 applied to this V_1. By altering the value V_1y to m_n/n we obtain a more general theorem of this sort:

THEOREM 12. *If $\phi=\phi(x)$ is irreducible over F, if*

$$f(x, y) = a_0(x)y^n + a_1(x)\phi^{m_1}y^{n-1} + a_2(x)\phi^{m_2}y^{n-2} + \cdots + a_n(x)\phi^{m_n},$$

where the $a_i(x)$ *are polynomials in* $F[x]$ *with* $a_0(x)\not\equiv 0 \pmod{\phi}$, $a_n(x)\not\equiv 0 \pmod{\phi}$, *and if the positive integers* m_i *are such that* m_n *is prime to* n *and* $n\cdot m_i \geqq i\cdot m_n$ *for* $i=1, \cdots, n$, *then* $f(x, y)$ *is irreducible in* $F[x, y]$, *except perhaps for a factor involving only* x.

A special case of this theorem, for $\phi(x)=x-\alpha$, was first stated by Königsberger.* If F is the field of all complex numbers the conditions of this theorem imply that the sheets of the Riemann surface hang together in a single cycle at $x=\alpha$, so that $f(x, y)$ is necessarily irreducible. In the theorem above, this single cycle is obtained by the first approximant; that is, by the first step of the Puiseux expansion. Similar theorems involving further steps of the expansion can be stated, but they are all consequences of our general criteria. The theorem corresponding to the point at infinity on the Riemann surface runs as follows:

THEOREM 13. *If* $a_i(x)$ *for* $i=0, 1, \cdots, n$ *are polynomials in* $F[x]$ *with no common factors except constants, if*

$$f(x, y) = a_0(x)y^n + a_1(x)y^{n-1} + \cdots + a_n(x),$$

if $\deg a_i(x)=\gamma_i$ *satisfies the conditions*

$$n(\gamma_i - \gamma_0) < i(\gamma_n - \gamma_0), \qquad \gamma_n - \gamma_0 \neq 0, \qquad (i = 0, 1, \cdots, n-1), \tag{13}$$

and if $\gamma_n-\gamma_0$ *is prime to* m, *then* $f(x, y)$ *is irreducible in* $F[x, y]$.

A related theorem stated by Perron† replaces (13) by the condition that there is an integer j such that

$$\gamma_j > \gamma_i, \qquad j\gamma_i < i\gamma_j \qquad (\text{for } i = 1, \cdots, n, \text{ but } i \neq j)$$

holds, while $a_0(x)=1$, and γ_j is prime to j. This theorem is apparently more general, but it can be deduced from Theorem 13 by simply interchanging x and y. The effect of the change can be visualized by constructing the Newton polygon for V_0.

Still other types of theorems are possible in case the coefficient field is not algebraically closed.

THEOREM 14. *If* $\phi(x)$ *is irreducible over* F, *if*

$$\psi(x, y) = y^m + a_1(x)y^{m-1} + \cdots + a_m(x), \qquad a_i(x) \text{ in } F[x],$$

is irreducible‡ $(\text{mod } \phi(x))$ *in* $F[x, y]$, *if*

---

* Königsberger, loc. cit., p. 63.

† O. Perron, *Neue Kriterien für die Irreduzibilität algebraischer Gleichungen*, Journal für die Mathematik, vol. 132 (1907), p. 304. See also Blumberg, loc. cit., p. 543.

‡ In other words if $\psi(\theta, y)$ is irreducible in $F(\theta)[y]$, where θ is a root of $\phi(x)=0$.

$$f(x, y) = \psi(x, y)^e + g(x, y)\phi(x)\psi(x, y) + r(x, y)\phi(x),$$

where the degrees of g and r in y satisfy $\deg_y r(x, y) < m$; $\deg_y g(x, y) < m(e-1)$, *and if* $r(x, y) \not\equiv 0 \pmod{\phi}$, *then* $f(x, y)$ *is irreducible in* $F[x, y]$.

Proof. If we omit the case where $m=1$ and $a_m(x) \equiv 0 \pmod{\phi}$, the conditions on $\psi(x, y)$ suffice to make ψ a key polynomial over the value

$$V_1 = [V_0F = 0, V_0x = 0, V_0\phi(x) = 1, V_1y = 0]$$

of $F[x, y]$. The only second stage approximant to f over V_0 is $V_2 = [V_1, V_2\psi(x, y) = 1/e]$. The given form of $f(x, y)$ shows that the expansion of f will satisfy the conditions of Theorem 2 for irreducibility. In the omitted case where $m=1$ and $a_m(x) \equiv 0 \pmod{\phi}$, the theorem is a direct consequence of Theorem 11.

Other theorems may be obtained by using the non-trivial values of the coefficient field F.

THEOREM 15. *If R is the field of rational numbers, if p is a prime, if* $\phi(x)$ *is a polynomial irreducible* (mod p), *if*

$$f(x, y) = y^n + g(x, y)\phi(x) + h(x, y)p,$$

where the polynomials g and h have integral coefficients and degrees in y less than n, and if either $g(x, 0) \not\equiv 0$ (modd $p, \phi(x)$) *or* $h(x, 0) \not\equiv 0$ (modd $p, \phi(x)$) *holds, then* $f(x, y)$ *is irreducible in* $R[x, y]$.

This theorem follows from Theorem 2 applied to the value

$$V_1 = [V_0p = 1, V_0x = 0, V_0\phi(x) = 1, V_1y = 1/n].$$

If $h=0$, Theorem 15 is a special case of Theorem 11. Other simple cases of Theorem 15 arise for $\phi(x)=x$. Under the same conditions on ϕ, g, and h, it is also possible to assert the irreducibility of

$$y^n + g(x, y)\phi(x)^e + h(x, y)p^f,$$

provided $(e, n) = 1 = (f, n)$.

Theorems on polynomials in several variables may also be stated. For example, the following theorem can be proven for any number of variables.

THEOREM 16. *If F is any field, if*

$$f(x, y, z) = a(y, z)x^k + b(x, z)y^m + c(x, y)z^n$$

where $a(y, z)$, $b(x, z)$, *and* $c(x, y)$ *are polynomials with coefficients in F, each with degrees less than k, m, and n in x, y, and z respectively, while* $a(0, 0) \neq 0$, $b(0, 0) \neq 0$, *and* $c(0, 0) \neq 0$, *and if* $(k, m) = 1$, $(k, n) = 1$, $(m, n) = 1$, *then* $f(x, y, z)$ *is irreducible in* $F[x, y, z]$.

Shanok* has also stated irreducibility criteria for polynomials in several variables in terms of convex polyhedra. Corresponding to each term $x^\alpha y^\beta p^\gamma$ in the polynomial there is a point with the coordinates (α, β, γ). Each side of the convex polyhedron on these points corresponds to a suitable absolute value $V=[Vp=\lambda, Vx=\mu, Vy=\nu]$, chosen so that three distinct terms in the original polynomial have the same minimum value. Thus, although Shanok's theorems are not special cases of ours, they could be stated in terms of absolute values.

---

* C. Shanok, *Convex polyhedra and criteria for irreducibility*, Duke Mathematical Journal, vol. 2 (1936), pp. 103–111.

Cornell University,
Ithaca, N. Y.

ON THE GROWTH OF ANALYTIC FUNCTIONS*

BY

NORMAN LEVINSON†

1. Pólya,‡ in a restricted case, and Bernstein,§ under rather general conditions, have, to state their results roughly, proved that the rate of growth of an analytic function along a line can be determined by its growth along a suitable sequence of discrete points on the line. In proving his results Bernstein uses certain rather deep theorems from the theory of Dirichlet series and points out¶ that as yet no proof of the results has been obtained using ordinary function theory.

Here we shall give a simple function-theoretic proof of a set of theorems which include those of Bernstein. We shall then refine our methods and obtain a set of new theorems which are remarkably precise. (See Theorem VII for example.)

We shall deal exclusively with functions which are analytic in a sector and of order 1.‖ For use here we can define the Phragmén-Lindelöf function for a function $f(z)$ analytic in a sector $|\text{am } z| \leqq \alpha$, as

$$h(\theta) = \limsup_{r\to\infty} \frac{\log |f(re^{i\theta})|}{r}, \qquad |\theta| \leqq \alpha. \tag{1.0}$$

The following theorems are among those which will be proved.

THEOREM I.** *Let $\phi(z)$ be analytic in some sector $|\text{am } z| \leqq \alpha$. Let $h(\theta)$ defined as in* (1.0) *be its Phragmén-Lindelöf function and let $h(0) = a$. Suppose*

$$h(\theta) \leqq a \cos\theta + b|\sin\theta|, \qquad |\theta| \leqq \alpha. \tag{1.1}$$

Let $\{z_n\}$ be a sequence of complex numbers such that

$$\lim_{n\to\infty} \frac{n}{z_n} = D, \tag{1.2}$$

* Presented to the Society, September 10, 1937; received by the editors March 4, 1937.

† National Research Fellow.

‡ G. Pólya, *Untersuchungen über Lücken und Singularitäten von Potenzreihen*, Mathematische Zeitschrift, vol. 29 (1924).

§ V. Bernstein, *Généralisation et Conséquences d'un théorème de Le Roy-Lindelöf*, Bulletin des Sciences Mathématiques, vol. 52 (1928); and *Séries de Dirichlet*, Chapter IX, Paris, 1933.

¶ V. Bernstein, *Séries de Dirichlet*, loc. cit., pp. 229 and 249.

‖ Functions of any other finite order can be transformed to functions of order 1 by $w = z^\rho$.

** In the case of real $\{z_n\}$ a proof of this theorem and a related gap theorem were given by the author in the Bulletin of the American Mathematical Society, vol. 42 (1936), p. 702.

where D is real, and such that for some $d>0$

$$|z_n - z_m| \geqq |n-m|\,d. \tag{1.3}$$

If

$$\pi D > b, \tag{1.4}$$

then

$$\limsup_{n\to\infty} \frac{\log|\phi(z_n)|}{|z_n|} = \limsup_{r\to\infty} \frac{\log|\phi(r)|}{r}. \tag{1.5}$$

In the case where the sequence $\{z_n\}$ is real, Theorem I is due to Bernstein.* We observe that (1.2) requires, if $z_n = r_n e^{i\theta_n}$, that $n/r_n \to D$ and that $|\theta_n| \to 0$ as $n\to\infty$. In other words all except a finite number of the z_n will lie in any sector containing the real axis.

A special case of Theorem I is the following:

THEOREM II. *Let* $\phi(z)$ *be analytic and of exponential type*† *in the half-plane* $|\text{am } z| \leqq \frac{1}{2}\pi$. *If L is defined by*

$$h(\tfrac{1}{2}\pi) + h(-\tfrac{1}{2}\pi) = 2\pi L, \tag{1.6}$$

and if $\{z_n\}$ *is a sequence satisfying* (1.2) *and* (1.3), *then*

$$D > L \tag{1.7}$$

implies (1.5).

It is easy to see by considering $\sin \pi z$ at the points $z_n = n$ that (1.7) is critical, for in this case $D = L = 1$ and it is clear that (1.5) is not true. Nevertheless we shall show that the condition (1.7) can be weakened considerably without disrupting the theorem. It is in weakening this condition that we obtain a new and very precise set of theorems.

The simplest of these theorems is

THEOREM VII. *Let* $\phi(z)$ *be analytic and of exponential type in the sector* $|\text{am } z| \leqq \frac{1}{2}\pi$. *Let*

$$\phi(iy) = O(1), \qquad |y| \to \infty.$$

Let $\{z_n\}$ *be a sequence of density*‡ $D \geqq 0$, *such that* $|\text{am } z_n| \to 0$ *as* $n\to\infty$ *and*

* Loc. cit., Chapter IX. Likewise Theorem II stated below is due to Bernstein in the case where the $\{z_n\}$ are real.

† A function $f(z)$ is of exponential type in a sector, $|\text{am } z| \leqq \alpha$, if $f(z) = O(e^{C|z|})$, $|\text{am } z| \leqq \alpha$, for some constant C.

‡ In case the density of $\{z_n\}$ is greater than zero, Theorem II gives Theorem VII at once. Thus this theorem is really of interest when $D=0$ in which case Theorem II cannot be applied.

$|z_n - z_m| \geqq |n-m|d$. *A necessary and sufficient condition that* (1.5) *hold is that*

$$\sum_1^\infty \frac{1}{|z_n|} = \infty. \tag{1.8}$$

Clearly Theorem VII is a much sharper result than that obtained by trying to apply Theorem II. For example, if $z_n = n \log (1+n)$, Theorem VII is applicable while Theorem II is not.

The condition $\phi(iy) = O(1)$ in Theorem VII can easily be replaced by

$$\int_{-\infty}^{\infty} \frac{\log^+ |\phi(iy)|}{1+y^2} dy < \infty.$$

Analogous results are true in Theorem VI which follows and in related theorems.

Another of these theorems is

THEOREM VI. *Let* $\phi(z)$ *be analytic and of exponential type in the half-plane* $|\text{am } z| \leqq \frac{1}{2}\pi$. *Let*

$$\phi(iy) = O(e^{\pi L|y|})$$

and let $\{\lambda_n\}$ *be an increasing sequence of positive numbers satisfying*

$$\lim_{n\to\infty} \frac{n}{\lambda_n} = D, \qquad \lambda_{n+1} - \lambda_n \geqq d > 0. \tag{1.9}$$

Let $\Lambda(u)$ *be the number of* $\lambda_n < u$. *If*

$$\int_1^\infty \frac{\Lambda(u) - uL}{u^2} du = \infty \tag{1.10}$$

*and if**

$$\Lambda(u) > Lu + C \tag{1.11}$$

for some C, *then*

$$\limsup_{n\to\infty} \frac{\log |\phi(\lambda_n)|}{\lambda_n} = \limsup_{x\to\infty} \frac{\log |\phi(x)|}{x}.$$

Theorem VI goes much further than Theorem II in that it *does not exclude the possibility of* $D = L$.

We shall first give simple function-theoretic proofs of Theorems I, II, and related results which are due to Bernstein in case $\{z_n\}$ is real. In §3 we

* This condition is by no means critical. It can for example be replaced by $\Lambda(u) > Lu - u^\alpha - C$, $\alpha < 1$. However condition (1.10) is the divergence condition (1.8) of Theorem VII, and, as in that theorem, it is easy to show that it is a best possible result.

will turn to the proofs of our new results such as Theorem VI, VII, and related theorems.

2. Here we concern ourselves with Theorems I and II and certain of their extensions.

The following result of Phragmén and Lindelöf will be of basic importance:*

THEOREM A. *Let $f(z)$ be an analytic function of $z=re^{i\theta}$, regular in the region R between two straight lines making an angle π/α at the origin, and on the lines themselves. Suppose $|f(z)|\leqq M$ on the lines and that as $r\to\infty$, $f(z)=O(e^{r^\beta})$ uniformly in R for some $\beta<\alpha$. Then $f(z)\leqq M$ throughout D.*

We also require

LEMMA 1. *Let $\{z_n\}$ satisfy* (1.2) *and* (1.3). *If*

$$F(z) = \prod_1^\infty\left(1-\frac{z^2}{z_n^2}\right), \tag{2.0}$$

then along the line am $z=\theta$

$$\lim_{r\to\infty}\frac{\log|F(re^{i\theta})|}{r} = \pi D|\sin\theta|, \qquad \theta\neq 0,\pi. \tag{2.1}$$

Also

$$\limsup_{|x|\to\infty}\frac{\log|F(x)|}{|x|} = 0. \tag{2.2}$$

Moreover

$$\frac{1}{F(z)} = O(e^{|z|\log|z|}), \qquad |z\pm z_n|\geqq \tfrac{1}{2}d, \tag{2.3}$$

and for any $\epsilon>0$

(2.4

$$\frac{1}{F'(z_n)} = O(e^{\epsilon|z_n|}).$$

For real $\{z_n\}$ all these results are well known.† (2.3) can be made much more precise but suffices for our purposes. The proof of this lemma is quite straightforward. Since it is very much the same as for real $\{z_n\}$ we omit it.

* See, e.g., Titchmarsh, *Theory of functions*, Oxford, 1932, p. 177.

† F. Carlson, *Über Potenzreihen mit endlich vielen verschiedenen Koeffizienten*, Mathematische Annalen, vol. 79 (1919), pp. 237–245, especially pp. 239–240.

Proof of Theorem I. We observe that there is no loss of generality in taking $\alpha<\frac{1}{2}\pi$. Clearly if (1.5) does not hold there exists a c such that

$$\limsup_{n\to\infty}\frac{\log|\phi(z_n)|}{|z_n|}<c<a=\limsup_{x\to\infty}\frac{\log|\phi(x)|}{x}. \tag{2.5}$$

It follows from (2.4) and (2.5) that

$$g(z)=\sum_1^\infty\frac{\phi(z_n)e^{-cz_n}}{F'(z_n)(z-z_n)}e^{cz}F(z) \tag{2.6}$$

is an entire function. Since $g(z_n)=\phi(z_n)$ it follows that

$$\psi(z)=\frac{\phi(z)-g(z)}{F(z)} \tag{2.7}$$

is analytic for $|\text{am } z|\leqq\alpha$. Using (1.1), (2.3), and (2.6), it follows that for $|z-z_n|\geqq\frac{1}{2}d$

$$\psi(z)=O(e^{|z|\log|z|}). \tag{2.8}$$

But $\psi(z)$ is analytic and (2.8) being true on the circles $|z-z_n|=\frac{1}{2}d$ must be true inside, since a function analytic in a domain takes its maximum value on the boundary of the domain. Therefore (2.8) holds in the entire sector.

From (1.1), (2.1), (2.6), and (2.7)

$$\psi(re^{\pm i\alpha})=O(\exp[r(a\cos\alpha+b\sin\alpha-\pi D\sin\alpha+\epsilon)]+\exp[cr\cos\alpha]),\ \epsilon>0.$$

Or setting $(\pi D-b)\tan\alpha=\gamma$ we have

$$\psi(re^{\pm i\alpha})=O(\exp[r\cos\alpha(a-\gamma+\epsilon\sec\alpha)]+\exp[cr\cos\alpha]). \tag{2.9}$$

Since $\pi D>b$, $\gamma>0$. If we take $\epsilon<\frac{1}{2}\gamma\cos\alpha$, then (2.9) becomes

$$\psi(re^{\pm i\alpha})=O(\exp[pr\cos\alpha]),\qquad p=\max(a-\tfrac{1}{2}\gamma,c).$$

In other words $\psi(z)e^{-pz}$ is bounded for am $z=\pm\alpha$. But by Theorem A, this and (2.8) implies that it is bounded in the entire sector $(-\alpha,\alpha)$. Thus in particular

$$\psi(x)=O(e^{px}),\qquad p=\max(a-\tfrac{1}{2}\gamma,c). \tag{2.10}$$

But by (2.7), $\phi(x)=\psi(x)F(x)+g(x)$. When we use (2.2), (2.6), and (2.10), this gives

$$\limsup_{x\to\infty}\frac{\log|\phi(x)|}{x}\leqq\max(a-\tfrac{1}{2}\gamma,c).$$

This contradicts the assumption that $h(0)=a$ and proves Theorem I.

Proof of Theorem II. There is no loss in generality in assuming that $h(\frac{1}{2}\pi)=h(-\frac{1}{2}\pi)=\pi L$.

Let us set

$$\limsup_{x\to\infty}\frac{\log|\phi(x)|}{x}=a.$$

Then clearly when we apply Theorem A to $\phi(z)\exp[(\pi L+\epsilon)iz-(a+\epsilon)z]$, $\epsilon>0$, in the upper right quadrant, it follows at once from Theorem A that here

$$h(\theta)\leqq a\cos\theta+\pi L|\sin\theta|.$$

Similarly this holds in the lower right quadrant. Theorem II now follows at once from Theorem I.

THEOREM III. *Let* $\Phi(z)$ *be an analytic function in the right half-plane* $|\operatorname{am} z|\leqq\frac{1}{2}\pi$ *such that*

(2.11) $$\Phi(re^{i\theta})=O(\exp[(a\log r\cos\theta+\pi b|\sin\theta|+\epsilon)r]),\qquad |\theta|\leqq\tfrac{1}{2}\pi,$$

where $a\geqq 0$, $b\geqq-\frac{1}{2}a$, *and* ϵ *is an arbitrary positive quantity. If* $\{z_n\}$ *is a sequence satisfying* (1.2) *and* (1.3) *and if*

(2.12) $$D>b+\tfrac{1}{2}a,$$

then

(2.13) $$\limsup_{n\to\infty}\frac{\log|\Phi(z_n)|}{|z_n|}\leqq\pi p$$

implies

(2.14) $$\Phi(re^{i\theta})=O(\exp[\pi r(p\cos\theta+b|\sin\theta|+\epsilon)]).$$

Proof. We shall assume that $a>0$, for if $a=0$ we have Theorem II. Let

(2.15) $$\phi(z)=\frac{\Phi(z)}{\Gamma(1+az)}\cdot$$

From Stirling's formula it follows easily that for $|\theta|\leqq\frac{1}{2}\pi$ and large r

$$\log|\Gamma(1+are^{i\theta})|-ar\log r\cos\theta+ar\theta\sin\theta+ar\cos\theta-\tfrac{1}{2}\log r=O(1).$$

Thus from (2.11)

(2.16) $$\phi(re^{i\theta})=O(\exp[(\pi b|\sin\theta|+a\theta\sin\theta+a\cos\theta+\epsilon)r]),\qquad |\theta|\leqq\tfrac{1}{2}\pi,$$

and from (2.13)

(2.17) $$\phi(z_n)=O(\exp[(-a\log|z_n|\cos\theta_n+\pi p+a+\epsilon)|z_n|]).$$

Using (2.12), (2.16), and (2.17) in Theorem II, we see that

$$\limsup_{r\to\infty} \frac{\log |\phi(re^{i\theta})|}{r} = -\infty, \qquad |\theta| < \tfrac{1}{2}\pi. \tag{2.18}$$

As in the proof of Theorem I we define

$$g(z) = \sum_1^\infty \frac{\phi(z_n)e^{Az_n}}{F'(z_n)(z-z_n)} e^{-Az}F(z), \tag{2.19}$$

where here A is any real number. We also consider $\psi(z) = \{\phi(z) - g(z)\}/F(z)$. As in Theorem I, $\psi(z)$ is an analytic function satisfying (2.8) for $|\text{am } z| \leqq \frac{1}{2}\pi$. Along the imaginary axis (assuming $\Re z_n > 1$ as we may with no restriction)

$$|\psi(iy)| \leqq \max \left| \frac{\phi(iy)}{F(iy)} \right| + \sum_1^\infty \left| \frac{\phi(z_n)e^{Az_n}}{F'(z_n)} \right|.$$

By (2.1), (2.12), and (2.16), $|\phi(iy)/F(iy)|$ is bounded. Thus there exists some $M_1 > 0$, which is entirely independent of A, such that

$$|\psi(iy)| \leqq M_1 + \sum_1^\infty \left| \frac{\phi(z_n)e^{Az_n}}{F'(z_n)} \right|. \tag{2.20}$$

Along am $z = \frac{1}{4}\pi i$ we have, using (2.1), (2.18), and (2.19),

$$\psi(re^{\pi i/4}) = O(\exp[-2Ar] + \exp[-Ar\cos\tfrac{1}{4}\pi]).$$

Using this and (2.20) in Theorem A, we see that $\psi(z)e^{Az}$ is bounded in the sectors $\frac{1}{4}\pi \leqq \text{am } z \leqq \frac{1}{2}\pi$ and $-\frac{1}{2}\pi \leqq \text{am } z \leqq \frac{1}{4}\pi$, or in the entire right half-plane. Again from Theorem A, (2.20) now implies that

$$|e^{Az}\psi(z)| \leqq M_1 + \sum_1^\infty \left| \frac{\phi(z_n)e^{Az_n}}{F'(z_n)} \right|, \qquad |\text{am } z| \leqq \tfrac{1}{2}\pi.$$

In particular then

$$|\psi(re^{\pi i/4})| \leqq \exp[-Ar\cos\tfrac{1}{4}\pi]\left(M_1 + \sum_1^\infty \left| \frac{\phi(z_n)e^{Az_n}}{F'(z_n)} \right|\right). \tag{2.21}$$

In the following we use M_2, M_3, etc., to represent positive constants independent of A. Using (2.19) and (2.1), we have

$$|g(re^{\pi i/4})| \leqq \exp[-Ar\cos\tfrac{1}{4}\pi + \pi Dr]M_2 \sum_1^\infty \left| \frac{\phi(z_n)e^{Az_n}}{F'(z_n)} \right|. \tag{2.22}$$

Since $\phi(z) = g(z) + \psi(z)F(z)$ we have, using (2.1), (2.21), and (2.22),

$$|\phi(re^{\pi i/4})| \leqq \exp[-Ar\cos\tfrac{1}{4}\pi + \pi Dr]M_3\left(1 + \sum_1^\infty \left| \frac{\phi(z_n)e^{Az_n}}{F'(z_n)} \right|\right).$$

Or setting $A = a \log r$, we have

$$\begin{aligned} &| \phi(re^{\pi i/4}) | \\ (2.23) \qquad &\leqq \exp[-ar\log r \cos \tfrac{1}{4}\pi + \pi D r] M_3 \left(1 + \sum_1^\infty \left| \frac{\phi(z_n)\exp[az_n \log r]}{F'(z_n)}\right|\right). \end{aligned}$$

If we use (2.4) and (2.17), we have

$$\begin{aligned} \sum_1^\infty &\left| \frac{\phi(z_n)\exp[az_n \log r]}{F'(z_n)}\right| \\ &= O\left(\sum_1^\infty \exp\left[-a|z_n|\log\frac{|z_n|}{r}\cos\theta_n + (\pi p + a + \epsilon)|z_n|\right]\right). \end{aligned}$$

But $\exp[-\lambda u \log(u/r)]$ has for its maximum value as u varies $e^{\lambda r/e}$. Thus if $B>1$ is so large that $\log B > 2(\pi p + a + \epsilon)/a$, then

$$\begin{aligned} \sum_1^\infty \left| \frac{\phi(z_n)\exp[az_n \log r]}{F'(z_n)}\right| &= O\Big((r \exp[(\pi p + a + \epsilon)Br + ar/e] \\ &\quad + \sum_{|z_n|>Br} \exp[-|z_n|(2\cos\theta_n - 1)(\pi p + a + \epsilon)]\Big) \\ &= O(\exp[Br(2a + \pi p + \epsilon)]). \end{aligned}$$

Using this in (2.23) gives

$$\phi(re^{\pi i/4}) = O(\exp[-ar\log r\cos\tfrac{1}{4}\pi + Cr]),$$

where $C = \pi D + B(2a + \pi p + \epsilon)$. Or in (2.15)

$$\Phi(re^{\pi i/4}) = O(e^{Cr}).$$

But this and (2.11) for $\theta = \pm\frac{1}{2}\pi$, used in Theorem A, show that $\Phi(z)$ is of exponential type in the right half-plane. Theorem III now follows at once from Theorem II.

THEOREM IV. *If in Theorem* III, (2.13) *is replaced by*

$$(2.24) \qquad \Phi(z_n) = O(|\exp[-kz_n \log|z_n|]|), \qquad k > 0,$$

then (2.14) *is replaced by*

$$(2.25) \quad \Phi(re^{i\theta}) = O(\exp[(-k\log r\cos\theta + \pi b|\sin\theta| + \epsilon)r]), \quad |\theta| \leqq \tfrac{1}{2}\pi.$$

Proof. This proof is identical with that of Theorem III except that in (2.17) we take account of (2.24) and replace a by $a+k$. This modification will then cause a corresponding change in (2.23) where a is again replaced by $a+k$ since now we set $A = (a+k)\log r$. This finally gives us

$$\Phi(re^{\pi i/4}) = O(\exp[-kr\log r\cos\tfrac{1}{4}\pi + Cr]).$$

From this we see that $\Phi(z)\Gamma(1+kz)$ is of exponential type and therefore Theorem II can be applied to obtain Theorem IV.

THEOREM V. *If $\Phi(z)$ satisfies the requirements of Theorem* III *with* (2.13) *replaced by* (2.24) *and*

$$k > 2b, \tag{2.26}$$

then $\Phi(z)=0$.

In proving this and subsequent results of this type we use a fundamental theorem* of Carleman, or rather a consequence of this theorem.

Let $f(z)$ be analytic in the half-plane $|\text{am } z| \leqq \frac{1}{2}\pi$ and let $R>1$. Then

$$\begin{aligned}\frac{1}{2\pi}\int_1^R\left(\frac{1}{y^2}-\frac{1}{R^2}\right)\log|f(iy)f(-iy)|\,dy \\ +\frac{1}{\pi R}\int_{-\pi/2}^{\pi/2}\log|f(Re^{i\theta})|\cos\theta\,d\theta + A > 0,\end{aligned} \tag{2.27}$$

where A is some number depending only on $f(z)$.

Proof. Applying (2.27) to $\Phi(z)$ we have, assuming it is not identically zero,

$$\begin{aligned}-A < \frac{1}{2\pi}\int_1^R\left(\frac{1}{y^2}-\frac{1}{R^2}\right)\log^+|\Phi(iy)\Phi(-iy)|\,dy \\ +\frac{1}{\pi R}\int_{-\pi/2}^{\pi/2}\log^+|\Phi(re^{i\theta})|\cos\theta\,d\theta.\end{aligned}$$

Using (2.25) and replacing A by another constant A_1, we get

$$-A_1 < (b+\epsilon)\log R - \tfrac{1}{2}k\log R,$$

or $2(b+\epsilon)\geqq k$. But by (2.26) this is impossible for arbitrarily small ϵ. Thus $\Phi(z)=0$.

3. In this section we consider Theorem VI and related theorems. We recall that Theorem VI may be valid even when $D=L$. The proof is quite different from those of the previous section.

The method of this section can best be presented by first using it to give an alternative proof of Theorem II.

Alternative proof of Theorem II. There is no restriction in assuming that

$$h(\tfrac{1}{2}\pi) = h(-\tfrac{1}{2}\pi) = \pi L. \tag{3.0}$$

* E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932, p. 130.

Moreover it is clear that (1.5) follows if we prove that

$$\limsup_{n\to\infty} \frac{\log |\phi(z_n)|}{|z_n|} \leqq 0 \tag{3.1}$$

implies that $h(0) \leqq 0$.

As in (2.6)

$$g(z) = \sum_{1}^{\infty} \frac{\phi(z_n)e^{-\epsilon z_n}}{F'(z_n)(z - z_n)} e^{\epsilon z}F(z) \tag{3.2}$$

for any $\epsilon > 0$, is an entire function of exponential type. And as in (2.7)

$$\psi(z) = \frac{\phi(z) - g(z)}{F(z)} \tag{3.3}$$

is analytic for $|\text{am } z| \leqq \frac{1}{2}\pi$. As in (2.8)

$$\psi(z) = O(\exp [|z| \log |z|]), \qquad |\text{am } z| \leqq \tfrac{1}{2}\pi. \tag{3.4}$$

Since $\phi(z)$ and $g(z)$ are of exponential type in $|\text{am } z| \leqq \frac{1}{2}\pi$, (2.1) and (3.3) give

$$\psi(re^{\pi i/4}) = O(e^{Br}) \tag{3.5}$$

for some $B > 0$. From (3.3) we have

$$|\psi(iy)| \leqq \left| \frac{\phi(iy)}{F(iy)} \right| + \sum_{1}^{\infty} \left| \frac{\phi(z_n)e^{-\epsilon z_n}}{F'(z_n)(iy - z_n)} \right|. \tag{3.6}$$

Since $D > L$, (3.0) and (2.1) imply that

$$\phi(iy)/F(iy) = O(\exp [-\tfrac{1}{2}\pi(D - L)|y|]). \tag{3.7}$$

From (3.6) and (3.7)

$$\psi(iy) = O\left(\frac{1}{|y|}\right). \tag{3.8}$$

Using (3.5) and (3.8), $ze^{-2Bz}\psi(z)$ is bounded along the imaginary axis and the line am $z = \frac{1}{4}\pi$. By Theorem A this and (3.4) implies that it is bounded in the entire right half-plane. Thus.

$$\frac{\psi(z)e^{-2Bz}}{1 + z} = O\left(\frac{1}{|z|^2}\right). \tag{3.9}$$

When we use this, it follows at once from the Cauchy integral theorem that for $x > 0$,

$$\frac{\psi(z)e^{-2Bz}}{1+z} = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{\psi(s)e^{-2Bs}}{1+s}\,\frac{ds}{z-s}$$
$$= \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{\psi(s)e^{-2Bs}}{1+s}\,ds \int_0^\infty e^{-u(z-s)}du$$
$$= \frac{1}{2\pi i}\int_0^\infty e^{-uz}du \int_{-i\infty}^{i\infty} \frac{\psi(s)e^{-2Bs}}{1+s}\,e^{us}ds.$$

Or if

$$H(u) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{\psi(s)e^{-2Bs}}{1+s}\,e^{us}ds, \tag{3.10}$$

then for $x>0$,

$$\frac{\psi(z)e^{-2Bz}}{1+z} = \int_0^\infty H(u)e^{-uz}du. \tag{3.11}$$

When we use (3.9) and close the path of integration to the right in (3.10), it is clear that

$$H(u) = 0, \qquad u < 0. \tag{3.12}$$

On the other hand, when we use (3.3) in (3.10), it follows that

$$\begin{aligned} H(u) &= \frac{1}{2\pi}\int_{-i\infty}^{i\infty} \frac{\phi(it)e^{-2iBt}}{F(it)(1+it)}\,e^{iut}dt \\ &\quad - \sum_1^\infty \frac{\phi(z_n)e^{-\epsilon z_n}}{F'(z_n)}\,\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{e^{s(u-2B+\epsilon)}}{(1+s)(s-z_n)}\,ds. \end{aligned} \tag{3.13}$$

Or for $u<2B-\epsilon$,

$$H(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\phi(it)e^{-2Bit}}{F(it)(1+it)}\,e^{iut}dt - \sum_1^\infty \frac{\phi(z_n)e^{z_n(u-2B)}}{F'(z_n)(1+z_n)}\,. \tag{3.14}$$

Clearly by (3.1), (2.4), and (1.2) the infinite series on the right of (3.14) represents an analytic function for $u<2B$. Again by (3.7) it follows that the infinite integral on the right of (3.14) represents an analytic function in u for $-\infty<u<\infty$. Since the sum of two analytic functions is analytic it follows that $H(u)$ is analytic for $u<2B-\epsilon$. But by (3.12), $H(u)=0$, $u<0$. Therefore $H(u)=0$, $u<2B-\epsilon$. Using this in (3.11), we have

$$\frac{\psi(x)e^{-2Bx}}{1+x} = \int_{2B-\epsilon}^\infty e^{-ux}H(u)du.$$

By (3.9) and (3.10), $H(u)$ is bounded. Thus

$$\frac{\psi(x)e^{-2Bx}}{1+x} = O(e^{-x(2B-\epsilon)})$$

or $\psi(x) = O(e^{2\epsilon x})$. If we recall that $\phi(x) = g(x) + F(x)\psi(x)$, it follows that

$$h(0) = \limsup_{x\to\infty} \frac{\log |\phi(x)|}{x} \leqq 2\epsilon.$$

Since ϵ is arbitrary, $h(0) \leqq 0$. This completes the proof.

The difference between this proof and those of §2 is that here we get a representation of $\psi(z)$ in terms of $H(u)$. In this section we are attempting to refine Theorem II so that $D > L$, (1.7), is not necessary. Let us see how such a change would affect the argument in the preceding theorem. It is clear that the crucial point in this argument is the paragraph following (3.14) and it is only with this that we need concern ourselves here.

It is convenient to write (3.14) as

$$H(u) = H_1(u) - H_2(u), \qquad u < 2B - \epsilon, \tag{3.15}$$

where

$$H_1(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\phi(it)e^{-2iBt}}{F(it)(1+it)} e^{iut}dt \tag{3.16}$$

and

$$H_2(u) = \sum_1^{\infty} \frac{\phi(z_n)e^{z_n(u-2B)}}{F'(z_n)(1+z_n)}. \tag{3.17}$$

It is clear that $H_2(u)$ is analytic for $u < 2B$ irrespective of how D compares with L. Therefore it is only with $H_1(u)$ that we need be concerned in changing (1.7). If $D = L$, then there need exist no $\delta > 0$ such that

$$\phi(it)/F(it) = O(e^{-\delta|t|})$$

and $H_1(u)$ need no longer be analytic.

Is there any weaker condition than analyticity on $H_1(u)$ that tells us that if $H_1(u) = H_2(u)$, $u < 0$, and $H_2(u)$ is analytic for $u < a$, then $H_1(u) = H_2(u)$, $u < a$? That there are such weaker conditions is shown by the following result.

THEOREM B. *If*

$$f(u) = \int_{-\infty}^{\infty} G(t)e^{itu}dt, \tag{3.18}$$

where $G(t) \epsilon L(-\infty, \infty)$ *and if for* $t > 0$

$$G(t) = O(e^{-\theta(t)}), \tag{3.19}$$

where $\theta(t)$ *is a monotone non-decreasing function such that*

$$\int_1^\infty \frac{\theta(t)}{t^2}\,dt = \infty, \tag{3.20}$$

then if $f(u)$ *coincides with an analytic function over some interval it coincides with the analytic function over its entire interval of analyticity on the* u *axis.**

In order to use this theorem with $H_1(u)$ for $f(u)$ it must be shown that $\phi(it)/F(it)$ satisfies a condition of the type (3.19) under the hypothesis of Theorem VI. This is done by

LEMMA 2. *If*

$$F(z) = \prod_1^\infty \left(1 - \frac{z^2}{\lambda_n^2}\right) \tag{3.21}$$

and if $\{\lambda_n\}$ *satisfies the requirements of Theorem* VI, *then there exists a non-decreasing function* $\theta(y)$, $y>0$, *such that for sufficiently large* $|y|$

$$e^{\pi L|y|}/F(iy) = O(|y|^{2C}e^{-\theta(|y|)}), \tag{3.22}$$

and such that

$$\int_1^\infty \frac{\theta(y)}{y^2}\,dy = \infty. \tag{3.23}$$

Proof. We assume that $\lambda_1 \geqq 1$ (since we can discard any λ_n's which are less than one). Clearly

$$\log |F(iy)| = \int_0^\infty d\Lambda(u) \log\left(1 + \frac{y^2}{u^2}\right)$$

$$= 2\int_0^\infty \frac{\Lambda(u)}{u}\,\frac{y^2}{y^2+u^2}\,du.$$

Since

$$2\int_0^\infty \frac{y^2}{y^2+u^2}\,du = \pi|y|,$$

we have

* In *On a class of non-vanishing functions*, Proceedings of the London Mathematical Society, vol. 41 (1936), p. 393, it is shown that if $f(u)$ vanishes over any interval it vanishes identically. Theorem B is closely related to this result. The proof of Theorem B will appear shortly in the Journal for Mathematics and Physics.

$$\log |F(iy)| - \pi L|y| = 2\int_0^\infty \frac{\Lambda(u) - Lu}{u}\, \frac{y^2}{y^2+u^2}\, du$$

$$= 4\int_0^\infty du\, \frac{y^2u}{(y^2+u^2)^2}\int_0^u \frac{\Lambda(u) - Lu}{u}\, du$$

$$= 4\int_0^\infty \frac{y^2u}{(y^2+u^2)^2}\, du \int_1^u (\Lambda(v) - Lv + C)\, \frac{dv}{v} - 4L\int_0^\infty \frac{y^2u}{(y^2+u^2)^2}\, du$$

$$- 4C\int_0^\infty \frac{y^2u\log u}{(y^2+u^2)^2}\, du$$

$$\geqq 4\int_{|y|}^\infty \frac{y^2u}{(y^2+u^2)^2}\, du \int_1^{|y|} (\Lambda(v) - Lv + C)\, \frac{dv}{v} - 4(L+C) - 2C\log|y|$$

$$= \int_1^{|y|} \frac{\Lambda(v) - Lv + C}{v}\, dv - 2C\log|y| + \text{constant}.$$

If we set

$$\theta(y) = \int_1^y \frac{\Lambda(v) - Lv + C}{v}\, dv, \qquad y > 1,$$

then by (1.11), $\theta(y)$ is non-decreasing. Also

$$\int_1^\infty \frac{\theta(y)}{y^2}\, dy = \int_1^\infty \frac{dy}{y^2}\int_1^y \frac{\Lambda(v) - Lv + C}{v}\, dv$$

$$= \int_1^\infty \frac{\Lambda(v) - Lv + C}{v^2}\, dv = \infty.$$

This proves the lemma.

Proof of Theorem VI. Here, as in the alternative proof of Theorem II, it suffices to show that $\limsup \log|\phi(\lambda_n)|/\lambda_n \leqq 0$ implies $h(0) \leqq 0$. The proof proceeds almost exactly like that of the alternative proof of Theorem II up to (3.14) except that we concern ourselves with $e^{-2Bz}\psi(z)/(1+z)^{2C+2}$ here. So in place of (3.14) we get

$$H(u) = \frac{1}{2\pi}\int_{-i\infty}^{i\infty} \frac{\phi(iy)e^{-2iBy}}{F(iy)(1+iy)^{2C+2}}\, e^{iuy}dy - \sum_1^\infty \frac{\phi(\lambda_n)e^{\lambda_n(u-2B)}}{F'(\lambda_n)(1+\lambda_n)^{2C+2}},$$

$$u < 2B - \epsilon.$$

As before the infinite series on the right is analytic for $u < 2B$ and $H(u) = 0$, $u < 0$, or by (3.15), $H_1(u) = H_2(u)$, $u < 0$. By Lemma 2

$$H_1(u) = \frac{1}{2\pi}\int_{-i\infty}^{i\infty} \frac{\phi(iy)e^{-2iBy}}{F(iy)(1+iy)^{2C+2}}\, e^{iuy}dy$$

satisfies the requirements of Theorem B. Since

$$H_1(u) = H_2(u) = \sum_1^\infty \frac{\phi(\lambda_n)e^{\lambda_n(u-2B)}}{F'(\lambda_n)(1+\lambda_n)^{2C+2}}, \qquad u < 0,$$

Theorem B implies that this is true for $u<2B-\epsilon$. That is $H(u)=0$ for $u<2B-\epsilon$. As in the alternative proof of Theorem II, this leads at once to $\psi(x)=O(e^{2\epsilon x})$. $h(0)\leqq 0$ follows at once completing the proof.

There are variations of Theorem VI such as the following:

THEOREM VI-A. *Theorem* VI *remains true if* (1.10) *is replaced by*

$$\phi(iy) = O(\exp[\pi L|y| - \theta(|y|)]),$$

where $\theta(y)$, $y>0$, *is a non-decreasing function of y which satisfies* (3.20).

The proof of this result is quite obvious from what precedes.

Before proving Theorem VII the following result very much like Lemma 2 is necessary.

LEMMA 3. *Let* $F(z)$ *be defined as in* (2.0) *with* $\{z_n\}$ *satisfying the requirements of Theorem* VII. *Then*

$$\frac{1}{F(iy)} = O(e^{-\theta(|y|)}),$$

where $\theta(y)$, $y>0$, *is a non-decreasing function of y satisfying* (3.20).

If $z_n=r_ne^{i\theta_n}$, there is no loss of generality in assuming that $|\theta_n|<\frac{1}{4}\pi$. We then have, if $\Lambda(u)$ is the number of $|z_n|<u$,

$$\begin{aligned}\log|F(iy)| &= \frac{1}{2}\sum_1^\infty \log\left(1+\frac{2y^2}{r_n^2}\cos 2\theta_n+\frac{y^4}{r_n^4}\right)\\ &\geqq \frac{1}{2}\sum_1^\infty \log\left(1+\frac{y^4}{r_n^4}\right) = \frac{1}{2}\int_0^\infty d\Lambda(u)\log\left(1+\frac{y^4}{u^4}\right)\\ &> \frac{\log 2}{2}\int_0^{|y|} d\Lambda(u) > \frac{1}{4}\Lambda(|y|).\end{aligned}$$

If we take $\theta(y)=\Lambda(|y|)$, then $|F(iy)|^{-1}=O(\exp[-\frac{1}{4}\theta(|y|)])$. To prove (3.20) we have

$$\int_0^\infty \frac{\Lambda(y)}{y^2}dy = \lim_{A\to\infty}\left[\int_0^A \frac{d\Lambda(y)}{y} - \frac{\Lambda(A)}{A}\right].$$

By (1.8)

$$\int_0^\infty \frac{d\Lambda(y)}{y} = \infty$$

and therefore (3.20) is satisfied.

Proof of Theorem VII. To show that (1.8) is sufficient we proceed exactly, as in the alternative proof of Theorem II up to (3.14). We then use Lemma 3 to apply Theorem B to $H_1(u)$, (3.16), and show that $H(u)=0$, $u<2B-\epsilon$. From this the fact that (3.1) implies $h(0)\leqq 0$ follows at once.

We now turn to the necessity of condition (1.8). Let us assume that

$$\sum_1^\infty \frac{1}{|z_n|} < \infty .$$

Then we shall show that no such result as (1.5) holds. Let

$$\phi(z) = \prod_1^\infty \frac{(z - z_n)(z - \bar{z}_n)}{(z + z_n)(z + \bar{z}_n)} .$$

Then $\phi(z)$ is analytic in the right half-plane. Moreover

$$|\phi(re^{i\theta})| = \prod_1^\infty \left| \frac{re^{i\theta} - z_n}{re^{i\theta} + z_n} \right| \left| \frac{re^{i\theta} - \bar{z}_n}{re^{i\theta} + \bar{z}_n} \right| \leqq 1, \qquad |\theta| \leqq \tfrac{1}{2}\pi .$$

Therefore $\phi(z)$ satisfies the requirements of Theorem VII. Moreover

$$\limsup_{n\to\infty} \frac{\log |\phi(z_n)|}{|z_n|} = -\infty .$$

If (1.5) could be applied it would give

$$\limsup_{x\to\infty} \frac{\log |\phi(x)|}{x} = -\infty .$$

Applying Theorem A to $e^{Cz}\phi(z)$, $C>0$, in the upper and lower right quadrants we see that it is bounded. Then again applying Theorem A to the bounded function $e^{Cz}\phi(z)$ in the right half-plane, we see that for $|\text{am } z| \leqq \frac{1}{2}\pi$, $|e^{Cz}\phi(z)| \leqq 1$. Since C can be made arbitrarily large this means that $\phi(z)=0$, which obviously is not the case. Thus (1.8) is a necessary condition in order that Theorem VII be true.

The analogue of Theorem III in this section is the following:

THEOREM VIII. *Let* $\Phi(z)$ *be an analytic function in the right half-plane* $|\text{am } z| \leqq \frac{1}{2}\pi$ *such that for any* $\epsilon>0$

$$\Phi(re^{i\theta}) = O(\exp\left[(a \log r \cos\theta + \epsilon\cos\theta + b|\sin\theta|)r\right]), \qquad |\theta| \leqq \tfrac{1}{2}\pi,$$

where $a \geqq 0$, $b \geqq -\frac{1}{2}a$, *and let* $\{\lambda_n\}$ *be a positive increasing sequence satisfying* (1.9). *Let* $\Lambda(u)$ *be the number of* $\lambda_n < u$. *If*

$$\int_1^\infty \frac{\Lambda(u) - (b + \frac{1}{2}a)u}{u^2} du = \infty,$$

if for some C, $\Lambda(u) > (b + \frac{1}{2}a)u - C$, *and*

$$\limsup_{n\to\infty} \frac{\log |\Phi(\lambda_n)|}{\lambda_n} \leqq \pi p, \tag{3.24}$$

then

$$\Phi(re^{i\theta}) = O(\exp[\pi r(p\cos\theta + b|\sin\theta| + \epsilon\cos\theta)]), \qquad |\theta| \leqq \tfrac{1}{2}\pi, \tag{3.25}$$

where ϵ *is an arbitrary positive quantity.*

Proof. As in Theorem III we consider $\phi(z) = \Phi(z)/\Gamma(1+az)$. By Theorem VI it follows that (2.18) holds. $g(z)$ is defined as in (2.19) and $\psi(z) = \{\phi(z) - g(z)\}/F(z)$.

In formulas analogous to (2.20), (2.21), and so on, we consider $\psi(z)/(1+z)^{2C}$ rather than just $\psi(z)$ as in Theorem III. Otherwise the proof now proceeds in precisely the same way as in Theorem III.

THEOREM IX. *If in Theorem* VIII (3.24) *is replaced by*

$$\Phi(\lambda_n) = O(\exp[-k\lambda_n \log |\lambda_n|]), \qquad k > 0,$$

then (3.25) *is replaced by*

$$\Phi(re^{i\theta}) = O(\exp[(-k\log r\cos\theta + \epsilon\cos\theta + \pi b|\sin\theta|)r]).$$

Proof. This theorem is related to Theorem VIII in the same way as Theorem IV is related to Theorem III and its proof follows almost at once from that of Theorem VIII just as that of Theorem IV follows almost at once from Theorem III.

THEOREM X. *Let* $\Phi(z)$ *satisfy the requirements of Theorem* VIII *with* (3.24) *replaced by*

$$\limsup_{n\to\infty} \frac{\log|\Phi(\lambda_n)| + 2b\lambda_n \log\lambda_n}{\lambda_n} = -\infty. \tag{3.26}$$

Then $\Phi(z) = 0$.

Proof. Applying Theorem IX to $e^{Bz}\Phi(z)$ for any $B > 0$, we have

$$|\Phi(z)| \leqq B_1(\exp[(-2b\log r\cos\theta - B\cos\theta + \pi b|\sin\theta|)r]), \qquad |\theta| \leqq \tfrac{1}{2}\pi, \tag{3.27}$$

where B_1 is a constant depending on B. Applying (2.27) to $\Phi(z)$ and using (3.27) we have for some A_1 depending only on $\Phi(z)$ (if $\Phi(z)$ does not vanish identically),

$$-A_1 \leqq \frac{1}{2\pi}\int_1^R \left(\frac{1}{y^2} - \frac{1}{R^2}\right)(2\pi b y)dy + \frac{1}{\pi R}\int_{-\pi/2}^{\pi/2} \log^+ |\Phi(Re^{i\theta})| \cos\theta \, d\theta.$$

Again using (3.27)

$$-A_1 - b\log R \leqq -\frac{2b\log R + B}{\pi}\int_{-\pi/2}^{\pi/2} \cos^2\theta \, d\theta + b\int_{-\pi/2}^{\pi/2} \cos\theta |\sin\theta| \, d\theta$$
$$+ \frac{\log B_1}{\pi R}\int_{-\pi/2}^{\pi/2} \cos\theta \, d\theta.$$

Or

$$-A_1 \leqq -\tfrac{1}{2}B + b + \frac{2}{\pi R}\log B_1.$$

Letting $R\to\infty$ we see that by choosing $B > 2A_1 + 2b$ we obtain a contradiction.

PRINCETON UNIVERSITY
AND
INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

ON DIFFERENTIAL GEOMETRY IN THE LARGE, I (MINKOWSKI'S PROBLEM)*

BY

HANS LEWY

Introduction. Hermann Minkowski,† in a fundamental paper on convex bodies, proposed the following Problem (M): to determine a convex, three-dimensional body B whose surface admits of a given Gaussian curvature $K(n)>0$, assigned as a continuous function of the direction of the interior normal n to the surface. After having stated three obviously necessary conditions for the function $K(n)$, Minkowski proceeds to solve an analogous Problem (M′) for polyhedra. He then considers a passage to the limit among the solutions of problems (M′) approximating (M), and establishes their convergence to a convex body B_0. This construction leaves open the question as to whether B_0 is a solution of (M). Minkowski remarks that B, if it exists, is, to within a translation, the uniquely determined solution of a certain third Problem (M″) of the calculus of variations and that B_0, too, is a solution of (M″).

Now if we assume (H): the surface of B_0 is differentiable to a sufficiently high order, then B_0 solves (M). However, Minkowski does not discuss (H), but proves instead that the mixed volume $V(B_0, B_0, C)$ of B_0 with an arbitrary convex body C may be computed as though B_0 were a solution of (M) and that, furthermore, a convex body is, except for a translation, uniquely determined by its mixed volume with the totality of convex bodies. While later authors have modified Minkowski's methods, there has been no improvement of his results as far as the hypothesis (H) is concerned.

Thus Minkowski's results are open to the same criticism that could be raised against the early solutions of Plateau's problem: namely, that instead of solving the proposed problem, a more general problem is treated whose solution coincides with that of the former only if the latter solution satisfies certain highly restrictive conditions, and no indication is presented that these conditions are actually satisfied.

The present paper contains a solution of (M) for the case of analytic $K(n)$. It does not involve the Brunn-Minkowski inequalities, nor, indeed, the idea of mixed volume. It uses instead the author's results on elliptic and

* Presented to the Society, November 28, 1936; received by the editors March 1, 1937.

† Minkowski, *Werke*, pp. 231–276, entitled *Volumen und Oberfläche*. See especially §10. Also Bonnesen-Fenchel, *Theorie der konvexen Körper*, Ergebnisse, vol. 3, no. 1 (1934), which contains a bibliography of related papers.

analytic equations of the Monge-Ampère type and establishes the existence of an analytic B by a continuity method.

In order to make this treatment of (M) complete, a new proof of the uniqueness of B is included. This proof is based on a modification of a beautiful idea of Cohn-Vossen* who, in a similar problem, reduced the uniqueness problem to the determination of a certain topological index of a vector field. Our modification is such as to allow an application of this idea to a wider class of related uniqueness problems in the large.

1. Let $u(x, y)$ be a homogeneous polynomial solution of Laplace's equation of degree $n>2$. A suitable rotation of the (x, y)-system transforms $u(x, y)$ into a constant multiple of $\Re[(x+iy)^n]$, and the asymptotic directions of the surface $u=u(x, y)$, determined by

$$u_{xx}dx^2 + 2u_{xy}dxdy + u_{yy}dy^2 = 0, \tag{1}$$

undergo the same rotation as the coordinates. Assuming this rotation effected, we obtain for the asymptotic directions

$$\Re[(x + iy)^{n-2}(d(x + iy))^2] = 0,$$

or, in polar coordinates r, θ,

$$\Re[r^{n-2}e^{i(n-2)\theta}(dx + idy)^2] = 0.$$

Thus the vector $dx+idy$ of the asymptotic direction forms the angle $-(n-2)\theta/2$ or $-(n-2)\theta/2+\pi/2$ with the x-axis; the two asymptotic directions are perpendicular, and either direction turns through an angle $-(n-2)\pi$ as we follow it along a Jordan curve containing the origin in its interior. In other words: the asymptotic directions form two distinct fields of directions with the origin as a singular point of index $-(n-2)/2$. We notice furthermore that the discriminant of (1) is a constant multiple of r^{2n-4} and vanishes only for $r=0$.

2. We prove now the following lemma:

LEMMA. *Let $F(x, y, u, p, q, r, s, t)$ be analytic in the neighborhood of $(x_0, y_0, u_0, p_0, q_0, r_0, s_0, t)$ and $4(\partial F/\partial r)(\partial F/\partial t)-(\partial F/\partial s)^2>0$. Let $u(x, y)$ and its first derivatives p, q and second derivatives r, s, t be a solution of $F=0$, analytic in a neighborhood of (x_0, y_0) and such that $u(x_0, y_0)=u_0$, $p(x_0, y_0)=p_0, \cdots, t(x_0, y_0)=t_0$. Assume $u'(x, y)$ and its derivatives p', q', r', s', t' to be a second analytic solution of $F=0$, coinciding together with its derivatives p', q', r', s', t', with u and its derivatives at the point (x_0, y_0). Then the difference $U=u-u'$ represents a surface $U(x, y)$ whose Gaussian curvature is negative in a sufficiently small*

* Cohn-Vossen, *Zwei Sätze über die Starrheit der Eiflächen*, Göttinger Nachrichten, 1927, pp. 125–134.

neighborhood of (x_0, y_0), with the exception of (x_0, y_0) itself, and the index of either of its distinct asymptotic directions is negative at (x_0, y_0), unless $u(x, y)$ is identically equal to $u'(x, y)$.

In terms of the second derivatives R, S, T of $U(x, y)$ the asympototic directions of $U=U(x, y)$ are given by

$$Q \equiv Rdx^2 + 2Sdxdy + Tdy^2 = 0; \tag{2}$$

and (2) admits of real solutions dx, dy if $RT-S^2<0$. To show that $RT-S^2<0$ express $F(x, y, u(x, y), p(x, y), \cdots)-F(x, y, u'(x, y), p'(x, y), \cdots)$ as a power series in $(x-x_0, y-y_0)$ and observe that the terms of lowest degree are given by

$$E \equiv a\overline{R} + b\overline{S} + c\overline{T},$$

where $\overline{R}$, $\overline{S}$, $\overline{T}$ are the second derivatives of the non-vanishing terms $\overline{U}$ of lowest degree $n>2$ in the development of U, and a, b, c are the values of $\partial F/\partial r$, $\partial F/\partial s$, $\partial F/\partial t$ at (x_0, y_0), for which by hypothesis $4ab-c^2>0$. Since $F(\cdots, u(x, y), \cdots)-F(\cdots, u'(x, y), \cdots)=0$, we have $E=0$, and this is possible only for $\overline{R}\overline{T}-\overline{S}^2<0$ or $\overline{R}=\overline{S}=\overline{T}=0$. Now $\overline{R}$, $\overline{S}$, $\overline{T}$ vanish simultaneously only at (x_0, y_0). For there exists a suitable linear transformation of the (x, y)-plane with determinant 1 which leaves (x_0, y_0) invariant and transforms $E=0$ into Laplace's equation. $\overline{U}$ is thereby transformed into a harmonic homogeneous polynomial of degree $n>2$ in $(x-x_0, y-y_0)$, and for such polynomials we proved in §1 that the discriminant of the second derivatives vanishes only at $(0, 0)$. Hence the development of $RT-S^2$ starts with the negative term $\overline{R}\overline{T}-\overline{S}^2$ and we conclude that $RT-S^2$ is negative for sufficiently small $|x-x_0|$, $|y-y_0|$ and vanishes only at $x=x_0$, $y=y_0$.

Since the linear transformation does not change the index of a field of directions, we conclude from §1 that the index of the field

$$\overline{Q} \equiv \overline{R}dx^2 + 2\overline{S}dxdy + \overline{T}dy^2 = 0 \tag{3}$$

is negative. Since, for sufficiently small values of $|x-x_0|$, $|y-y_0|$, the directions of the field (2) differ arbitrarily little from those of (3), the index of (3) is negative at (x_0, y_0). This completes the proof of the Lemma.

3. We may now prove the following:

THEOREM 1. *Two closed convex analytic surfaces S and S' are congruent if they possess the same positive Gaussian curvature K at points for which their inner normals are parallel and similarly directed.*

By parallel normals we map S and S' on the unit sphere σ. An arbitrary equator divides σ and, since the map is one-to-one, S and S', into two regions

in each of which the sphere as well as S and S' assume the form $Z=Z(X, Y)$ in suitably chosen rectangular coordinates. Upon introducing

$$x = \frac{\partial Z}{\partial X}, \qquad y = \frac{\partial Z}{\partial Y} \tag{4}$$

as independent variables instead of (X, Y) and setting

$$H(x, y) = - Z(x, y) + xX(x, y) + yY(x, y), \tag{5}$$

we obtain

$$X = H_x, \qquad Y = H_y, \qquad Z = - H + xX + yY. \tag{6}$$

How S and S' satisfy the condition that their curvatures for corresponding parallel normals, i.e., for the same value of (x, y), are the same positive function $K(x, y)$, whence S and S' are solutions of

$$\left(1 + \left(\frac{\partial Z}{\partial X}\right)^2 + \left(\frac{\partial Z}{\partial Y}\right)^2\right)^2 K(x, y) = \frac{\partial^2 Z}{\partial X^2}\,\frac{\partial^2 Z}{\partial Y^2} - \left(\frac{\partial^2 Z}{\partial X \partial Y}\right)^2$$

or

$$(1 + x^2 + y^2)^2 K(x, y) = \frac{\partial(x, y)}{\partial(X, Y)},$$

or, finally, of

$$H_{xx}H_{yy} - H_{xy}^2 = K^{-1}(x, y)(1 + x^2 + y^2)^{-2}. \tag{7}$$

It is readily shown that the second fundamental form of the surface (6) is

$$(H_{xx}dx^2 + 2H_{xy}dxdy + H_{yy}dy^2)(1 + x^2 + y^2)^{-1/2}. \tag{8}$$

Suppose that the formulas (4) and (5) lead to a function $H(x, y)$ if applied to S, and to $H'(x, y)$ if applied to S'. Then the statement of Theorem 1 is equivalent to the equation

$$H'(x, y) = H(x, y) + l(x, y),$$

where $l(x, y)$ is a linear function of (x, y).

Consider, with Cohn-Vossen, the congruence points of S and S', i.e., those points for which their normals and their second fundamental forms coincide. If all of their points were congruence points, we should have

$$H_{xx} = H'_{xx}, \qquad H_{xy} = H'_{xy}, \qquad H_{yy} = H'_{yy},$$

and the theorem is proved. In the alternative case we shall show the existence of at least one congruence point. The two second differential forms of S and S' may both be assumed to be positive definite. The equation

$$(H_{xx} - H'_{xx})dx^2 + 2(H_{xy} - H'_{xy})dxdy + (H_{yy} - H'_{yy})dy^2 = 0 \tag{9}$$

obtained by setting the two forms equal to one another, determines two distinct directions tangential to σ at each point of σ which is not a congruence point. For the ellipses of the (dx, dy)-plane

$$H_{xx}dx^2 + 2H_{xy}dxdy + H_{yy}dy^2 = 1$$

and

$$H'_{xx}dx^2 + 2H'_{xy}dxdy + H'_{yy}dy^2 = 1$$

have, by (7), the same area $\pi(1+x^2+y^2)K^{1/2}(x, y)$. As they are concentric but not identical, they intersect in four distinct points; the ratios of their coordinates are the two distinct solutions $dx: dy$ of (9).

It is impossible to construct a field of tangential directions on the sphere without singularities, and the sum of the indices of these singularities equals 2 if there are only finitely many singularities. Now choose, at an arbitrary point of σ, one of the two directions (9) and extend, by continuity, this choice over the whole of σ. If there were no congruence points, we should obtain a field of tangential directions on σ without singularities. Hence there is at least one congruence point (x_0, y_0).

Subtracting if necessary a linear function from $H'(x, y)$ we may assume that, at (x_0, y_0), $H(x, y)$ and $H'(x, y)$ coincide with their derivatives up to the second order, without affecting the truth of (7) nor that of (9). But now our lemma implies that unless H and H' are identical, the congruence point (x_0, y_0) is isolated and has a negative index. Summing over all indices of all singularities we still obtain a negative number in contradiction to the general fact mentioned above. Hence S and S' are congruent.

THEOREM 1'. *Consider a sequence of closed convex analytic surfaces $S(\tau)$ of positive curvature such that their corresponding functions $H(x, y, \tau)$ depend analytically on (x, y, τ) for small values of τ. Suppose that all surfaces $S(\tau)$ have a common point of contact corresponding to the same point (x, y) of σ. Assume that for each point of $S(0)$ the derivative $(\partial K/\partial\tau)_{\tau=0}$ vanishes. Then we have also $(\partial H/\partial\tau)_{\tau=0}=0$.*

Abbreviate the operator $(\partial/\partial\tau)_{\tau=0}$ by the use of the symbol δ and consider the field of tangential directions dx, dy on σ, determined by

$$\delta H_{xx}dx^2 + 2\delta H_{xy}dxdy + \delta H_{yy}dy^2 = 0. \tag{9'}$$

Unless $\delta H_{xx}=\delta H_{xy}=\delta H_{yy}=0$, there are, at each point (x, y), two real and distinct directions satisfying (9'). For we deduce from (7) by differentiation

$$H_{yy}\delta H_{xx} + H_{xx}\delta H_{yy} - 2H_{xy}\delta H_{xy} = 0, \tag{7'}$$

which yields $\delta H_{xx}\delta H_{yy}-(\delta H_{xy})^2<0$ in view of $H_{xx}H_{yy}-H_{xy}^2>0$.

As in the preceding proof we conclude first the existence of a point P on σ for which $\delta H_{xx} = \delta H_{yy} = \delta H_{xy} = 0$. Assume this not to be true identically.

Since an addition of a linear function of (x, y) to δH does not affect the truth of (7′) and (9′), we may assume that the field of directions determined by (9′) in the neighborhood of P corresponds to a function δH that vanishes at P together with its derivatives up to the second order. This function δH may be considered as the difference between the following two solutions of (7′): δH itself and the identically vanishing solution. From the lemma we see that P is an isolated singularity of the field defined by (9′) and that the index of P is negative. The proof of Theorem 1′ can now easily be completed by recalling the conclusion at the end of the proof of Theorem 1.

4. We state now the following:

DEFINITION. *Let $d\omega$ be the surface element of the unit sphere σ and let ξ, η, ζ be the cosines of the angle of its normal with the axes of the (X, Y, Z)-system. A function F of a point of σ is called admissible if it depends analytically on the point of σ and if*

$$\int F\xi d\omega = \int F\eta d\omega = \int F\zeta d\omega = 0. \tag{10}$$

THEOREM 2. *For every admissible positive function K on σ there exists a closed convex analytic surface S whose curvature, considered as function of the interior normal vector* (*of length* 1), *equals K.*

First we shall demonstrate the following statement:

II. *Assume that for small values of a parameter τ an admissible positive function K depends analytically on the point of σ and τ, and that for $\tau = 0$ there exists a surface $S(0)$ with $K(0)$ as corresponding curvature function. Then there exists an analytic closed convex surface $S(\tau)$ with $K(\tau)$ as corresponding curvature provided $|\tau|$ is sufficiently small.*

Let M be an arbitrary but analytic function on the unit sphere σ. Continue its definition from the sphere into the three-dimensional space containing it by assuming M to be zero at the center and linear on every ray issuing from the center. M thereby becomes a homogeneous function of degree 1 in any system of rectangular coordinates (ξ, η, ζ) with the origin as center and M is analytic everywhere except at the origin. We have from the homogeneity

$$M = M_\xi \xi + M_\eta \eta + M_\zeta \zeta. \tag{11}$$

Introduce

$$x = -\xi/\zeta, \qquad y = -\eta/\zeta, \qquad \zeta H(x, y) = -M. \tag{12}$$

We obtain

$$H_x = M_\xi, \qquad H_y = M_\eta. \tag{13}$$

On σ the coordinates (x, y) may be used for either hemisphere $\zeta<0$ or $\zeta>0$. From (13) we conclude that H_x, H_y are analytic on the whole of σ if we define them on the equator $\zeta=0$ by continuity.

We find

$$\begin{aligned} \pm(\xi, \eta, \zeta) &= \left(\frac{x, y, -1}{(1+x^2+y^2)^{1/2}}\right), \\ d\omega &= \frac{dxdy}{(1+x^2+y^2)^{3/2}}. \end{aligned} \tag{14}$$

Whenever M is such that $\partial(H_x, H_y)/\partial(x, y)\neq 0$, the reasoning of p. 260 shows that

$$\frac{\partial(H_x, H_y)}{\partial(x, y)}\cdot(1+x^2+y^2)^2 = \phi(H, H) \tag{15}$$

is the reciprocal Gaussian curvature of the surface

$$X = M_\xi, \qquad Y = M_\eta, \qquad Z = M_\zeta, \tag{16}$$

whence we conclude that ϕ remains invariant under the change which x, y, and H undergo as ξ, η, ζ are subjected to an orthogonal transformation that is sufficiently near the identity. This fact is obviously independent of the condition $\partial(H_x, H_y)/\partial(x, y)\neq 0$.

The integral over an arbitrary region

$$\mathfrak{J} = \iint \mp (H_{xx}H_{yy} - H_{xy}^2)dxdy = \iint \phi(H, H)\zeta d\omega$$

can be transformed into the integral over its boundary

$$\frac{1}{2}\int (H_x dH_y + H_y dH_x).$$

As the expressions H_x, H_y are analytic on the unit sphere we can extend the integration in $\mathfrak{J}$ over the whole of σ and obtain $\mathfrak{J}=0$ since the boundary integrals over the equator, generated by the integration over the upper and lower hemispheres, annul each other.

Using the invariance of $\phi(H, H)$, we obtain similarly

$$\iint \phi(H, H)\xi d\omega = \iint \phi(H, H)\eta d\omega = \iint \phi(H, H)\zeta d\omega = 0. \tag{17}$$

We also find for the associated bilinear form

$$\phi(H_1, H_2) = (1 + x^2 + y^2)^2(H_{1xx}H_{2yy} + H_{1yy}H_{2xx} - 2H_{1xy}H_{2xy})$$

that

$$\iint \phi(H_1, H_2)\xi d\omega = \iint \phi(H_1, H_2)\eta d\omega = \iint \phi(H_1, H_2)\zeta d\omega = 0. \tag{18}$$

In other words, $\phi(H_1, H_2)$ is admissible if M_1 and M_2 are analytic on the sphere.

After these preliminary remarks we return to the given function $K^{-1}(\tau)$ and develop it into a power series in τ,

$$K^{-1}(\tau) = \kappa(\tau) = \kappa_0 + \tau\kappa_1 + \tau^2\kappa_2 + \cdots . \tag{19}$$

We try to determine a function $M(\xi, \eta, \zeta; \tau)$, homogeneous of first degree in ξ, η, ζ,

$$M(\tau) = M_0 + M_1\tau + M_2\tau^2 + \cdots , \tag{20}$$

depending analytically on ξ, η, ζ, τ (excepting the point $\xi=\eta=\zeta=0$) and being in the following relation to $K^{-1}(\tau)$: If we introduce by (12) the quantities $x, y, H(x, y; \tau)$, then (7) holds for all sufficiently small values of $|\tau|$. Let us develop the equation (7)

$$\phi(H, H) = \kappa$$

into a power series in τ and set the coefficients of both members equal. Denoting again by the operator δ the differentiation with respect to τ at $\tau=0$, we find the equations:

$$\delta^n M = \delta^n H(x, y)/(1 + x^2 + y^2)^{1/2}, \quad M_0 = H(x, y)/(1 + x^2 + y^2)^{1/2}, \tag{21}$$

$$H_{xx}\delta H_{yy} + H_{yy}\delta H_{xx} - 2H_{xy}\delta H_{xy} = \kappa_1(x, y)(1 + x^2 + y^2)^{-2},$$

$$H_{xx}\delta^2 H_{yy} + H_{yy}\delta^2 H_{xx} - 2H_{xy}\delta^2 H_{xy} = 2!\kappa_2(x, y)(1 + x^2 + y^2)^{-2} - 2(\delta H_{xx}\delta H_{yy} - \delta H_{xy}\delta H_{xy}), \tag{22}$$

. . . .

Since by hypothesis $K^{-1}(\tau)$ is admissible for all τ in question, the same is true for the coefficients:

$$\iint \kappa_\nu \xi d\omega = \iint \kappa_\nu \eta d\omega = \iint \kappa_\nu \zeta d\omega = 0, \qquad (\nu = 0, 1, 2, \cdots).$$

Suppose that we have found $\delta^n M(\xi, \eta, \zeta)$ for $n<m$, in accordance with (22), and analytic on the sphere. Then the preliminary remarks show that the right-hand member in the mth equation yields zero if we multiply it in turn by

$$(1+x^2+y^2)^2\xi d\omega, \qquad (1+x^2+y^2)^2\eta d\omega, \qquad (1+x^2+y^2)^2\zeta d\omega$$

and integrate over the whole sphere. We therefore are entitled to apply a theorem of Hilbert* in which he states the now well known alternative for elliptic differential equations on the sphere for the special case of the equation with admissible right-hand member r,

$$(1+x^2+y^2)^2[H_{xx}(\delta H)_{yy}+H_{yy}(\delta H)_{xx}-2H_{xy}(\delta H)_{xy}]=r,$$
$$\delta H\cdot(1+x^2+y^2)^{-1/2}=\delta M,$$

and establishes the existence of a solution δM, which is differentiable infinitely many times as the coefficients of the differential equation are. Applying Hilbert's result to the mth equation (22) we find the function $\delta^m M$. Since this equation is elliptic because $S(0)$ is convex and accordingly $H_{xx}H_{yy}-H_{xy}^2>0$, we conclude from the analyticity of the equation the analyticity of $\delta^m M$ on the sphere.

It remains to be seen that for sufficiently small $|\tau|$ the series

$$M=M_0+\sum_1^\infty\frac{\delta^n M}{n!}\tau^n$$

converges. This can be true only if we eliminate the arbitrariness that affects the determination of $\delta^m M$, since we may add to $\delta^m M$ an arbitrary linear combination of ξ, η, ζ with constant coefficients and still retain a solution of the differential equation for $\delta^m M$. Denote by $L[v]$ the linear elliptic differential expression in v on the sphere which in coordinates (x, y) reduces to

$$L(v)\equiv(1+x^2+y^2)^2(H_{xx}u_{yy}+H_{yy}u_{xx}-2u_{xy}H_{xy}),\quad u=v(1+x^2+y^2)^{1/2}.$$

Since $L[v]=0$ has precisely the three linearly independent solutions $v=\xi, \eta, \zeta$, $L[v]$ admits of a Green's function of the second kind $G(A;B)$, where A and B are two points of the sphere. With the aid of one such $G(A;B)$ we solve the equations (22), written in the abbreviated form

$$L[\delta^\nu M]=f^\nu, \qquad (\nu=1,2,\cdots),$$

by setting

$$\delta^\nu M=\iint G(A;B)f^\nu(B)d\omega_B, \tag{23}$$

and thereby disposing of the arbitrariness in the determination of $\delta^\nu M$.

We shall say "a function f on the sphere satisfies a Hölder condition of ex-

* D. Hilbert, *Grundzüge einer Allgemeinen Theorie der Linearen Integralgleichungen*, Leipzig and Berlin, 1912, p. 250.

ponent α and coefficient C" if for arbitrary distinct points Q_1 and Q_2 of spherical distance $\overline{Q_1Q_2}>0$ we have

$$|f(Q_1) - f(Q_2)| \leqq C\overline{Q_1Q_2}^{\alpha},$$

where α is a positive constant less than 1. Similarly "the derivatives of f satisfy a Hölder condition of exponent α and coefficient C" if $0<\alpha<1$ and for arbitrary $Q_1 \neq Q_2$

$$|f_s(Q_1) - f_s(Q_2)| \leqq C\overline{Q_1Q_2}^{\alpha}, \qquad |f_n(Q_1) - f_n(Q_2)| \leqq C\overline{Q_1Q_2}^{\alpha},$$

where f_s is the derivative in the direction of the great circle joining Q_1 to Q_2 and f_n is the derivative in the direction normal to this circle. Similarly for higher derivatives.

With these notations the familiar estimates of the theory of linear elliptic differential equations are applied to the integral $v(A) = \int\int G(A; B)f(B)d\omega_B$ and lead to the following:

LEMMA. *If f is bounded by C and satisfies a Hölder condition of exponent α and coefficient C, then $v(A) = \int\int G(A; B)f(B)d\omega_B$ and its first and second derivatives are bounded by hC and satisfy a Hölder condition of exponent α and coefficient Ch where h does not depend on f.*

From the convergence of the series (19), with the use of the Heine-Borel theorem, we derive the existence of a number $\bar{\rho}>0$ such that

$$|\kappa_\nu| < \frac{1}{\bar{\rho}^\nu}, \qquad |\kappa_\nu'| < \frac{1}{\bar{\rho}^\nu}, \qquad (\nu = 1, 2, \cdots),$$

where κ_ν' stands for the directional derivative of κ_ν at an arbitrary point in an arbitrary direction. Hence there also exists a number $\rho>0$ such that $|\kappa_\nu| < 1/\rho^\nu$ and κ_ν satisfies a Hölder condition of a certain exponent α and coefficient $1/\rho^\nu$.

The first equation (22) shows that δM and its first and second derivatives are bounded by h/ρ and satisfy a Hölder condition of exponent α and coefficient h/ρ.

Consider the equation

$$z^2(\tau) = \frac{1}{64h^2} - \sum_1^\infty \frac{\tau^n}{\rho^n}. \tag{24}$$

It admits of two roots for $z(\tau)$ and in particular the root

$$z(\tau) = -\frac{1}{8h} + \sum_1^\infty \frac{c_n\tau^n}{n!\rho^n}.$$

We have the following system of recurrent formulas from (24):

$$2z(0)z'(0) = -\frac{1}{\rho} \qquad \text{or} \qquad c_1 = \frac{4h}{\rho},$$

$$2z(0)z''(0) = -\frac{2!}{\rho^2} - 2z'^2(0) \qquad \text{or} \qquad c_2 = \frac{4h\cdot 2!}{\rho^2} + 8hc_1^2,$$

$$2z(0)z'''(0) = -\frac{3!}{\rho^3} - 6z'(0)z''(0) \qquad \text{or} \qquad c_3 = \frac{4h\cdot 3!}{\rho^3} + 24hc_1c_2,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

On the other hand consider the successive bounds and Hölder coefficients C_ν for $\delta^\nu M$ and its first and second derivatives as obtained from (22). We have evidently the same law by which to form the successive inequalities*

$$C_1 \leqq \frac{4h}{\rho},$$

$$C_2 \leqq 4h\cdot\frac{2!}{\rho^2} + 8hC_1^2,$$

$$C_3 \leqq 4h\cdot\frac{3!}{\rho^3} + 24hC_1C_2,$$

$$\cdots\cdots\cdots\cdots$$

As all terms involved are positive we conclude $C_\nu \leqq c_\nu$, $(\nu = 1, 2, \cdots)$. Hence the series

$$M(\tau) = M(0) + \sum_1^\infty \frac{\delta^n M}{n!}\tau^n$$

converges together with its first and second derivatives with respect to ξ, η, ζ uniformly for sufficiently small $|\tau|$ since $z(\tau)$ is a majorant series. In order to complete the proof of statement II we have to show that $M(\tau)$ depends analytically on ξ, η, ζ for all sufficiently small $|\tau|$ and $(\xi, \eta, \zeta) \neq (0, 0, 0)$.

Denoting by $M_n(\tau)$ the nth partial sum of $M(\tau)$, by $\kappa_n(\tau)$ that of $\kappa(\tau)$, and by $H_n(x, y; \tau)$ the function $M_n(\tau)(1+x^2+y^2)^{1/2}$, we have for arbitrary $\epsilon > 0$ and uniformly for all points of the sphere

$$|\Delta_n| \equiv |(H_{nxx}H_{nyy} - H_{nxy}^2)(1 + x^2 + y^2)^2 - \kappa_n(\tau)| < \epsilon,$$

provided n is large enough. For Δ_n is a polynomial whose term of lowest degree in τ is of degree n; if we replace everywhere in Δ_n the δM, $\delta^2 M$, $\cdots$ and

* Observe the invariance of $\phi(\delta H, \delta H)$.

their first and second derivatives with respect to (x, y) by their upper bounds C_ν and form the sum of the absolute values of all terms thus obtained, we obtain less than the nth remainder in the similarly formed majorant of

$$(H_{xx}(\tau)H_{yy}(\tau) - H^2_{xy}(\tau))(1 + x^2 + y^2)^2, \text{ where } H(\tau) = M(\tau)\cdot(1 + x^2 + y^2)^{1/2};$$

and this majorant converges since it is a polynomial of convergent power series in τ.

Thus $\lim_{n\to\infty} H_n(x, y; \tau) = H(x, y; \tau)$ is a solution of the analytic elliptic equation

$$(H_{xx}(\tau)H_{yy}(\tau) - H^2_{xy}(\tau))(1 + x^2 + y^2) = K^{-1}(\tau)$$

and its second derivatives satisfy a Hölder condition; hence, by a well known theorem,* $H(x, y; \tau)$ is analytic in x and y for every closed bounded region of the (x, y)-plane. This result may, of course, be formulated invariantly by stating that $M(\tau)$ depends analytically on the point of the sphere for small values of $|\tau|$.

It will be observed that the proof of statement II which is thereby completed follows precisely the routine way of solving a functional equation in the neighborhood of a value τ_0 of a parameter τ entering the equation, if it can be solved for $\tau = \tau_0$. Our theorem on Monge-Ampère equations† is, however, the essential tool which makes it possible to derive our present Theorem 2 from the statement II.

Returning to the hypotheses of Theorem 2 we form, with the given positive admissible distribution of reciprocal curvature on σ, the family of positive admissible distributions

$$K^{-1}(\tau) = 1 - \tau + \tau K^{-1},$$

which for $\tau = 0$ reduces to the reciprocal curvature of σ itself and for $\tau = 1$ to that of the surface to be determined. Let τ' be the greatest value of τ, $0 \leqq \tau' \leqq 1$, such that for every positive ϵ there exists an analytic surface $S(\tau)$ of curvature $K(\tau)$ with $\tau' - \epsilon < \tau < \tau'$. We shall show that $\tau' = 1$. First of all, by II, $\tau' > 0$. Since for all values of τ in $0 \leqq \tau \leqq 1$ the curvature $K(\tau)$ is bounded from below by a fixed positive number, a theorem of Bonnet‡ shows that all existing $S(\tau)$ have a diameter which is bounded from above. Now take an arbitrary normal of σ and introduce coordinates (ξ, η, ζ) such that its inter-

---

* E. Hopf, *Über den funktionalen, insbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung*, Mathematische Zeitschrift, vol. 34 (1931), pp. 194–233.

† Hans Lewy, *A priori limitations for solutions of elliptic Monge-Ampère equations*, II, these Transactions, vol. 41 (1937), pp. 365–374, especially Theorem 2′ on p. 374.

‡ Cf. Blaschke, *Differentialgeometrie*, vol. 1, Berlin, 1924, p. 150.

section with σ becomes (0, 0, 1). Our introduction of the (x, y)-system will give this point the coordinates (0, 0). Let $H(x, y; \tau)\ (1+x^2+y^2)^{-1/2}$ be the distance of the tangent plane of $S(\tau)$ from a fixed point for which we take the center of gravity of $S(\tau)$. Then we have $|H(x, y, \tau)| < 2\beta$ where β is the upper bound of the diameter of $S(\tau)$ and (x, y) is restricted to the circle $x^2+y^2<1$. Apply our theorem on Monge-Ampère equations to a sequence of solutions $H(x, y, \tau)$ of (7) for which the parameter τ tends to τ'. We obtain a subsequence converging to an analytic solution $H(x, y, \tau')$ of (7) with $\tau=\tau'$. Since the origin of the (x, y)-system corresponds to an arbitrary normal of σ, the Heine-Borel Lemma shows the existence of a closed analytic surface $S(\tau')$ of curvature $K(\tau')$. Now $S(\tau')$ may be made the starting point for the construction of $S(\tau)$ for infinitely many values of τ, greater than and close to τ', with the aid of II. Thus the assumption that τ' be less than 1 and at the same time the greatest value in every neighborhood of which there are smaller values of τ admitting a surface $S(\tau)$, has led to a contradiction. Hence $\tau'=1$ and $S(\tau')=S(1)$ exists.

University of California
Berkeley Calif.

ON THE SERIES FOR THE PARTITION FUNCTION*

BY
D. H. LEHMER

1. **Introduction.** In 1917 Hardy and Ramanujan† gave the following asymptotic formula for the number $p(n)$ of partitions of n,

$$p(n) = \frac{1}{\pi 8^{1/2}} \sum_{k=1}^{[\alpha n^{1/2}]} A_k(n) k^{1/2} \frac{d}{dn}\left(\frac{e^{c\lambda_1/k}}{\lambda_n}\right) + O(n^{-1/4}), \tag{1.1}$$

where

$$c = \pi(2/3)^{1/2}, \qquad \lambda_n = (n - 1/24)^{1/2}, \qquad \alpha > 0,$$

and the coefficients $A_k(n)$ are defined by

$$A_1(n) = 1, \qquad A_2(n) = (-1)^n, \qquad A_3(n) = 2\cos\left[\pi(12n-1)/18\right]$$

and in general

$$A_k(n) = \sum_{(\rho)} \omega_{\rho,k} e^{-2n\rho\pi i/k}, \tag{1.2}$$

where ρ ranges over those numbers which are less than k and prime to k. Here $\omega_{\rho,k}$ are certain $24k$th roots of unity which arise in the theory of modular functions and are defined by (1.4) and (1.5).

Without knowledge of the behavior of $A_k(n)$ for large values of k other than the obvious fact that

$$A_k(n) = O(k) \tag{1.3}$$

Hardy and Ramanujan were unable to decide several questions about (1.1). For instance, if α is given, (1.1) gives $p(n)$ to within half a unit for all sufficiently large n. Just how large n must be was not discovered. Whether (1.1) would converge if extended to infinity and what is the least number of terms that need be taken were other questions depending on the magnitude of $A_k(n)$.

Quite recently Rademacher‡ has shown that if in (1.1) we replace e^x by $2 \sinh x$ we obtain an infinite series for $p(n)$ (with $\alpha = \infty$) whose convergence follows easily from (1.3). This striking result enables one to estimate the

* Presented to the Society, March 27, 1937; received by the editors March 11, 1937.

† Proceedings of the London Mathematical Society, (2), vol. 17, pp. 75–115.

‡ Proceedings of the London Mathematical Society, (2), vol. 43, pp. 241–254.

difference between $p(n)$ and the first N terms of the series of Hardy and Ramanujan. This estimate, of course, depends on $A_k(n)$ so that information about the general behavior of $A_k(n)$ for large as well as small values of k is important in this connection.

Apart from these questions there is the problem of actually using (1.1) to determine isolated values of $p(n)$ for n large. The task of evaluating $A_k(n)$ from its definition is quite formidable when k is large. Hardy and Ramanujan gave $A_k(n)$ for $k \leqq 18$ as sums of cosines, while the actual values of $A_k(n)$ for $k \leqq 20$ and all n have been tabulated recently.* The apparent intricacy of $A_k(n)$ would seem definitely to restrict the usefulness of (1.1).

It would therefore seem desirable to make an intensive study of $A_k(n)$. In a recent paper† we have proved that the series (1.1) would diverge if extended to infinity. This result was obtained from a simple estimate of $A_k(n)$, where k is a square of a prime. In this paper we give formulas of $A_k(n)$ as a single term thus eliminating the necessity of any sort of tables of $A_k(n)$. This result enables us to give close estimates for $A_k(n)$, and to answer the questions mentioned above, and makes feasible the application of (1.1) to any number of terms.

The method employed in the present paper depends in part on showing that $A_k(n)$ may be transformed into "generalized Kloosterman sums." These in turn may be evaluated by a slight extension of the results of Salié.‡ The results of this paper were first obtained independently of Kloosterman sums. Considerable space is saved, however, by referring to results already published. Section 2 is devoted to multiplication theorems for $A_k(n)$ which reduce the evaluation of $A_k(n)$ to the case in which k is a prime or a power of a prime.§ In §3 these evaluations are carried out. The final section applies the results of the preceding sections to the Hardy-Ramanujan and Rademacher series.

The quantities $\omega_{\rho,k}$ appearing in the definition (1.2) of $A_k(n)$ are given by

$$\omega_{\rho,k} = \left(\frac{-\rho}{k}\right) \exp\left[-\left\{\frac{1}{4}(k-1)+\frac{1}{12}\left(k-\frac{1}{k}\right)(2\rho+\bar{\rho}-\rho^2\bar{\rho})\right\}\pi i\right] \tag{1.4}$$

if k is odd, and by

---

* Journal of the London Mathematical Society, vol. 11 (1936), pp. 117–118. Erratum: for $A_{20}(n)$ read $A_{20}(n+5)$.

† Journal of the London Mathematical Society, vol. 12, pp. 171–176.

‡ Mathematische Zeitschrift, vol. 34 (1931), pp. 91–109.

§ This multiplicative property, discovered empirically, could have been anticipated eight years ago from a result of Estermann: Hamburger Abhandlungen, vol. 7 (1929), p. 91.

$$\omega_{\rho,k}=\left(\frac{-k}{\rho}\right)\exp\left[-\left\{\frac{1}{4}(2-k\rho-\rho)+\frac{1}{12}\left(k-\frac{1}{k}\right)(2\rho+\bar{\rho}-\rho^2\bar{\rho})\right\}\pi i\right] \tag{1.5}$$

when k is even. Here (a/b) is the symbol of Jacobi and $\rho\bar{\rho}\equiv 1 \pmod{k}$.

If we substitute $\exp[\{1-(a/b)\}(\pi i/2)]$ for the symbol (a/b) in (1.4) and (1.5), we obtain after a few simple reductions the following expressions for $A_k(n)$:

$$A_k(n) = \sum_{(\rho)} \exp\left[f_n(\rho)\,\frac{\pi i}{12k}\right] \qquad (k \text{ odd}), \tag{1.6}$$

$$A_k(n) = \sum_{(\rho)} \exp\left[g_n(\rho)\,\frac{\pi i}{12k}\right] \qquad (k \text{ even}), \tag{1.7}$$

where ρ ranges over a complete system of residues prime to k and where the functions f and g are given by

$$f_n(\rho)=f_n(\rho,\,k)=-\left\{24n\rho+6k\left(\frac{-\rho}{k}\right)+3k(k-3)+(k^2-1)(2\rho+\bar{\rho}-\rho^2\bar{\rho})\right\}, \tag{1.8}$$

$$g_n(\rho)=g_n(\rho,\,k)=-\left\{24n\rho+6k\left(\frac{-k}{\rho}\right)-\rho(k+1)(k+2)-(k^2-1)(\rho^2-1)\bar{\rho}\right\}. \tag{1.9}$$

The number $1-24n$ plays the dominant role in what follows and is abbreviated by writing

$$\nu = 1 - 24n. \tag{1.10}$$

From (1.6) and (1.7) it is seen that $f(\rho)$ and $g(\rho)$ need only be determined modulo $24k$. The following fundamental congruences are used many times in the sequel and are set forth here for reference. If k is odd

$$f_n(\rho) \equiv \nu\rho + \bar{\rho} \pmod{k \text{ or } 3k}^* \tag{1.11}$$

according as 3 is prime to k or not.

$$f_n(\rho) \equiv 0 \pmod{3} \tag{1.12}$$

if k is prime to 3.

$$f_n(\rho) \equiv 2k\left(\frac{-\rho}{k}\right) + k - 3 \pmod{8}. \tag{1.13}$$

If $k=2^\lambda k_1$ where k_1 is odd,

$$g_n(\rho) \equiv \nu\rho + \bar{\rho} \pmod{k_1 \text{ or } 3k_1}^* \tag{1.14}$$

* In these congruences $\bar{\rho}$ stands for $1/\rho \pmod{M}$ where M is the modulus of the congruence.

according as k_1 is prime to 3 or not. If k_1 is prime to 3

$$g_n(\rho) \equiv 0 \pmod 3. \tag{1.15}$$

For every k_1

$$g_n(\rho) \equiv \nu\rho + \bar{\rho} + 2k\left(\frac{-k}{\rho}\right) + k(k+3)\rho \pmod{2^{\lambda+3}}.* \tag{1.16}$$

2. **Multiplication theorems.** We shall derive three theorems for expressing $A_k(n)$ as a product of two A's whose subscripts are coprime integers whose product is k. This enables us to evaluate $A_k(n)$ for all k when the values of A_q are known, where q runs over all powers of primes.

THEOREM 1. *If k_1 and k_2 are odd coprime integers, then*

$$A_{k_1}(n_1)A_{k_2}(n_2) = A_{k_1k_2}(n_3), \tag{2.1}$$

where $n_3 \equiv k_1^2 n_2 + k_2^2 n_1 - (k_1^2 + k_2^2 - 1)/24 \pmod{k_1k_2}$.

Remark. In case k_1 or k_2 is a multiple of 3, the numerator of the fraction $(k_1^2+k_2^2-1)/24$ is also a multiple of 3 and the fraction becomes of the form $M/8$. In any case, then, the quantity n_3 may be replaced by an integer modulo k_1k_2.

Proof. We consider first the product

$$A_{k_1}(n_1)A_{k_2}(n_2) = \sum_{(\rho_1)}\sum_{(\rho_2)} \exp\left[\left\{k_2 f_{n_1}(\rho_1, k_1) + k_1 f_{n_2}(\rho_2, k_2)\right\}\frac{\pi i}{12k_1k_2}\right], \tag{2.2}$$

where ρ_1 and ρ_2 range over the numbers less than and prime to k_1 and k_2 respectively. For each value of these summation indices we define ρ_3 by the system of congruences

$$\rho_3 \equiv \rho_1/k_2 \pmod{k_1}, \tag{2.3}$$

$$\rho_3 \equiv \rho_2/k_1 \pmod{k_2}. \tag{2.4}$$

It is clear that as ρ_1 and ρ_2 range over their respective values the numbers ρ_3 modulo k_1k_2 run over the numbers $<k_1k_2$ and prime to k_1k_2 so that

$$A_{k_1k_2}(n_3) = \sum_{(\rho_3)} \exp\left[f_{n_3}(\rho_3, k_1k_2)\frac{\pi i}{12k_1k_2}\right]. \tag{2.5}$$

We show that every term of (2.2) is equal to the corresponding term of (2.5), where the correspondence is determined by (2.3) and (2.4) and n_3 is defined by (2.1). This amounts to showing that

---

* See footnote on p. 273.

(2.6) $$D_1 = f_{n_3}(\rho_3, k_1k_2) - \{k_2f_{n_1}(\rho_1, k_1) + k_1f_{n_2}(\rho_2, k_2)\} \equiv 0 \pmod{24k_1k_2}.$$

In the first place if neither k_1 nor k_2 is divisible by 3, then it follows from (1.12) that

$$D_1 \equiv 0 \pmod{3}.$$

On the other hand we may suppose from symmetry that k_2 is divisible by 3 if not both k_1 and k_2 are prime to 3. Therefore let $k=3k_2$ or k_2 according as 3 does or does not divide k_1k_2. Our task is then to show that

$$D_1 \equiv 0 \pmod{8kk_1}.$$

We consider first the modulus k_1. By (1.11) we have

$$D_1 \equiv \nu_3\rho_3 + \bar{\rho}_3 - k_2(\nu_1\rho_1 + \bar{\rho}_1) \pmod{k_1},$$

where

(2.7) $$\nu_3 = 1 - 24n_3 = k_2^2\nu_1 + k_1^2\nu_2.$$

Hence in view of (2.3) we have

$$\nu_3\rho_3 + \bar{\rho}_3 \equiv k_2\nu_1\rho_1 + k_2\bar{\rho}_1 \pmod{k_1},$$

so that

$$D_1 \equiv 0 \pmod{k_1}.$$

If $k=k_2$, the same argument shows that $D_1 \equiv 0 \pmod{k}$. In case $k=3k_2$ we note that $f_{n_1}(\rho_1, k_1) \equiv 0 \pmod{3}$ so that

$$D_1 \equiv f_{n_3}(\rho_3, k_1k_2) - k_1f_{n_2}(\rho_2, k_2) \equiv \nu_3\rho_3 + \bar{\rho}_3 - k_1(\nu_2\rho_2 + \bar{\rho}_2) \pmod{k},$$

but from (2.7) $\nu_3 \equiv k_1^2\nu_2 \pmod{k}$, so that

$$D_1 \equiv k_1(k_1\rho_3 - \rho_2)\left(\nu_2 - \frac{1}{k_1\rho_2\rho_3}\right) \pmod{k}.$$

By (2.4) the second factor is a multiple of k_2 while the third factor is a multiple of 3 since $\nu_2 \equiv 1 \pmod{3}$, and $k_1\rho_2\rho_3 \equiv \rho_2^2 \pmod{k_2}$ so that $k_1\rho_2\rho_3 \equiv 1 \pmod{3}$. Hence

$$D_1 \equiv 0 \pmod{kk_1}.$$

There remains to show that $D_1 \equiv 0 \pmod{8}$. By (1.13) we have

(2.8) $$\begin{aligned} D_1 \equiv\ & 2k_1k_2\left(\frac{-\rho_3}{k_1k_2}\right) + k_1k_2 - 3 \\ & - \left\{2k_1k_2\left[\left(\frac{-\rho_1}{k_1}\right) + \left(\frac{-\rho_2}{k_2}\right)\right] + 2k_1k_2 - 3k_1 - 3k_2\right\} \pmod{8}, \end{aligned}$$

while, by (2.3) and (2.4),

$$(2.9)\qquad \begin{aligned}\left(\frac{-\rho_3}{k_1k_2}\right) &= \left(\frac{-\rho_3}{k_1}\right)\left(\frac{-\rho_3}{k_2}\right) = \left(\frac{-\rho_1}{k_1}\right)\left(\frac{k_2}{k_1}\right)\left(\frac{-\rho_2}{k_2}\right)\left(\frac{k_1}{k_2}\right) \\ &= \left(\frac{-\rho_1}{k_1}\right)\left(\frac{-\rho_2}{k_2}\right)(-1)^{(k_1-1)(k_2-1)/4}.\end{aligned}$$

We now separate two cases according as k_1 and k_2 are both of the form $4x-1$ or not. In the affirmative case we have

$$\left(\frac{-\rho_3}{k_1k_2}\right) = -\left(\frac{\rho_1}{k_1}\right)\left(\frac{\rho_2}{k_2}\right), \text{ and } k_1k_2 \equiv 1 \pmod 4,$$

so that (2.8) becomes

$$\begin{aligned} D_1 &\equiv 2\left\{-1+\left(\frac{\rho_1}{k_1}\right)+\left(\frac{\rho_2}{k_2}\right)-\left(\frac{\rho_1}{k_1}\right)\left(\frac{\rho_2}{k_2}\right)\right\} + (k_1+3)(k_2+3)+4 \\ &\equiv -2\left\{1-\left(\frac{\rho_1}{k_1}\right)\right\}\left\{1-\left(\frac{\rho_2}{k_2}\right)\right\} + 4 + 4 \equiv 0 \pmod 8.\end{aligned}$$

In case not both k_1 and k_2 are of the form $4x-1$ we have from (2.9)

$$\left(\frac{\rho_3}{k_1k_2}\right) = \left(\frac{-\rho_1}{k_1}\right)\left(\frac{-\rho_2}{k_2}\right)$$

and

$$\begin{aligned} D_1 &\equiv 2k_1k_2\left\{-1-\left(\frac{-\rho_1}{k_1}\right)-\left(\frac{-\rho_2}{k_2}\right)+\left(\frac{-\rho_1}{k_1}\right)\left(\frac{-\rho_2}{k_2}\right)+2\right\} \\ &\qquad - 4k_1k_2 + (k_1+3)(k_2+3)+4 \\ &\equiv 2k_1k_2\left\{1-\left(\frac{-\rho_1}{k_1}\right)\right\}\left\{1-\left(\frac{-\rho_2}{k_2}\right)\right\} \\ &\qquad - 4(k_1k_2-1) + (k_1+3)(k_2+3) \equiv (k_1+3)(k_2+3) \\ &\equiv 0 \pmod 8,\end{aligned}$$

since at least one factor is a multiple of 4.

This completes the proof of Theorem 1.

THEOREM 2. *Let k be odd and λ be an integer >1, then*

$$A_k(n_1)A_{2^\lambda}(n_2) = (-1)^{2^{\lambda-2}}A_{2^\lambda k}(n_3),$$

where

$$(2.10)\qquad n_3 \equiv k^2n_2 + 2^{2\lambda}n_1 - (k^2-1+2^{2\lambda})/24 \pmod{2^\lambda k}.$$

Proof. Since

$$(-1)^{2^{\lambda-2}} = \exp(3\cdot 2^{2\lambda}k\pi i/12\cdot 2^{\lambda}k),$$

we have to show that the product

$$(2.11) \qquad A_k(n_1)A_{2^\lambda}(n_2) = \sum_{(\rho_1)}\sum_{(\rho_2)} \exp\left[\{2^\lambda f_{n_1}(\rho_1, k) + kg_{n_2}(\rho_2, 2^\lambda)\}\pi i/12\cdot 2^\lambda k\right]$$

is equal to

$$(2.12) \quad (-1)^{2^{\lambda-2}}A_{2^\lambda k}(n_3) = \sum_{(\rho_3)} \exp\left[\{g_{n_3}(\rho_3, 2^\lambda k) + 3\cdot 2^{2\lambda}k\}\pi i/12\cdot 2^\lambda k\right].$$

In fact we show that, provided ρ_3 is related to ρ_1 and ρ_2 by means of the system

$$(2.13) \qquad \rho_3 \equiv \rho_1/2^\lambda \pmod{k},$$

$$(2.14) \qquad \rho_3 \equiv \rho_2/k \pmod{2^\lambda},$$

then the corresponding terms of (2.11) and (2.12) are equal. To this effect we consider the difference

$$D_2 = g_{n_3}(\rho_3, 2^\lambda k) + 3\cdot 2^{2\lambda}k - \left[2^\lambda f_{n_1}(\rho_1, k) + kg_{n_2}(\rho_2, 2^\lambda)\right]$$

and prove that it is divisible by $24\cdot 2^\lambda k$.

Consider first the modulus 3 if k is not divisible by 3. Then by (1.12) and (1.15) each term of D_2 is a multiple of 3. Next consider D_2 modulo k or $3k$ according as 3 is prime to k or not. Referring to (1.11) and (1.14) we find that

$$D_2 \equiv \nu_3\rho_3 + \bar\rho_3 - 2^\lambda(\nu_1\rho_1 + \bar\rho_1) \pmod{k \text{ or } 3k}.$$

But $\nu_3 = 1-24n_3 \equiv 2^{2\lambda} - 24\cdot 2^{2\lambda}n_1 \equiv 2^{2\lambda}\nu_1 \pmod{k \text{ or } 3k}$, so that

$$D_2 \equiv 2^\lambda(2^\lambda\rho_3 - \rho_1)\left(\nu_1 - \frac{1}{2^\lambda\rho_1\rho_3}\right) \pmod{k \text{ or } 3k}.$$

The second factor is a multiple of k by (2.13) and, in case 3 divides k, the third factor is also a multiple of 3. Hence

$$D_2 \equiv 0 \pmod{3k}.$$

We must now show that $D_2 \equiv 0 \pmod{2^{\lambda+3}}$. Using (1.13) and (1.16) we have

$$\begin{aligned} D_2 \equiv{} & \nu_3\rho_3 + 2^{\lambda+1}k\left(\frac{-2^\lambda k}{\rho_3}\right) + 2^{2\lambda}\rho_3 + 3\cdot 2^\lambda k\rho_3 + \bar\rho_3 + 3\cdot 2^{2\lambda}k \\ & - 2^{\lambda+1}k\left(\frac{-\rho_1}{k}\right) - 2^\lambda k + 3\cdot 2^\lambda \\ & - \left\{k\nu_2\rho_2 + 2^{\lambda+1}\left(\frac{-2^\lambda}{\rho_2}\right)k + 2^{2\lambda}k\rho_2 + 3\cdot 2^\lambda k\rho_2 + k\bar\rho_2\right\} \pmod{2^{\lambda+3}}. \end{aligned}$$

But since $\nu_3 \equiv k^2\nu_2 + 2^{2\lambda} \pmod{2^{\lambda+3}}$ we obtain on collecting terms

$$
\begin{aligned}
D_2 \equiv (k\rho_3 - \rho_2)\left(k\nu_2 + 2^{2\lambda}k - \frac{1}{\rho_2\rho_3}\right) + 2^{2\lambda}\rho_3 - 2^{\lambda}3k(\rho_2 - \rho_3) + 3\cdot 2^{2\lambda}k \\
(2.15) \qquad - 2^{\lambda}(k-3) + 2^{\lambda+1}k\left[\left(\frac{-2^{\lambda}k}{\rho_3}\right) - \left(\frac{-\rho_1}{k}\right) - \left(\frac{-2^{\lambda}}{\rho_2}\right)\right] \pmod{2^{\lambda+3}}.
\end{aligned}
$$

Consider for the moment the Jacobi symbols. We will show that the quantity

$$Q = \left(\frac{-2^{\lambda}k}{\rho_3}\right) - \left(\frac{-\rho_1}{k}\right) - \left(\frac{-2^{\lambda}}{\rho_2}\right) \equiv -(-1)^{(k_1-1)(\rho_3-1)/4} \pmod{4}.$$

We consider separately the cases λ even and λ odd.

If λ is even, we have by (2.14), $\rho_2 \equiv \rho_3 k \pmod 4$, and

$$\left(\frac{-2^{\lambda}k}{\rho_3}\right) = \left(\frac{-1}{\rho_3}\right)\left(\frac{k}{\rho_3}\right), \quad \left(\frac{-2^{\lambda}}{\rho_2}\right) = \left(\frac{-1}{\rho_2}\right) = \left(\frac{-1}{\rho_3}\right)\left(\frac{-1}{k}\right).$$

By (2.13)

$$\left(\frac{-\rho_1}{k}\right) = \left(\frac{-2^{\lambda}}{k}\right)\left(\frac{\rho_3}{k}\right) = \left(\frac{-1}{k}\right)\left(\frac{\rho_3}{k}\right).$$

Hence

$$
\begin{aligned}
Q &= \left(\frac{-1}{\rho_3}\right)\left(\frac{k}{\rho_3}\right) - \left(\frac{-1}{k}\right)\left(\frac{\rho_3}{k}\right) - \left(\frac{-1}{\rho_3}\right)\left(\frac{-1}{k}\right) \\
&= \left\{\left(\frac{k}{\rho_3}\right) - \left(\frac{-1}{k}\right)\right\}\left\{\left(\frac{-1}{\rho_3}\right) + \left(\frac{\rho_3}{k}\right)\right\} - \left(\frac{k}{\rho_3}\right)\left(\frac{\rho_3}{k}\right) \\
&\equiv -(-1)^{(k-1)(\rho_3-1)/4} \pmod 4.
\end{aligned}
$$

If λ is odd, we have $\lambda > 2$, and $\rho_2 \equiv \rho_3 k \pmod 8$ so that

$$\left(\frac{-2^{\lambda}k}{\rho_3}\right) = \left(\frac{-2}{\rho_3}\right)\left(\frac{k}{\rho_3}\right), \quad \left(\frac{-2^{\lambda}}{\rho_2}\right) = \left(\frac{-2}{\rho_3}\right)\left(\frac{-2}{k}\right)$$

and

$$\left(\frac{-\rho_1}{k}\right) = \left(\frac{-2}{k}\right)\left(\frac{\rho_3}{k}\right).$$

Hence, in this case,

$$
\begin{aligned}
Q &= \left\{\left(\frac{k}{\rho_3}\right) - \left(\frac{-2}{k}\right)\right\}\left\{\left(\frac{-2}{\rho_3}\right) + \left(\frac{\rho_3}{k}\right)\right\} - \left(\frac{\rho_3}{k}\right)\left(\frac{k}{\rho_3}\right) \\
&\equiv -(-1)^{(k-1)(\rho_3-1)/4} \pmod 4.
\end{aligned}
$$

With this result we return to (2.15). We note first that on account of (2.14) we have

$$k\rho_3 - \rho_2 = 2^\lambda h, \tag{2.16}$$

so that the factor

$$\left(k\nu_2 + 2^{2\lambda}k - \frac{1}{\rho_2\rho_3}\right)$$

of (2.15) may be considered modulo 8 only. Since $\nu_2 \equiv 1 \pmod 8$ we find that by (2.16) the first term of (2.15) is congruent to

$$\rho_3 h^2 \cdot 2^{2\lambda} \pmod{2^{\lambda+3}}.$$

Hence the first four terms of (2.15) become congruent to

$$\begin{aligned} &2^{2\lambda}\left[\rho_3(h^2 + 1) + 3k(h + 1)\right] - 3\cdot 2^\lambda k\rho_3(k - 1) \\ &\qquad \equiv -3\cdot 2^\lambda k\rho_3(k-1) \equiv 3\cdot 2^\lambda\rho_3(k-1) \pmod{2^{\lambda+3}}, \end{aligned}$$

since the quantity in square brackets is even and $2\lambda+1 \geqq \lambda+3$. Hence we have

$$\begin{aligned} D_2 &\equiv 2^\lambda\left[3\rho_3(k-1) - (k-3) + 2kQ\right] \\ &\equiv 2^\lambda\left[3\rho_3 k - 3\rho_3 - k + 3 - 2k\cdot(-1)^{(k-1)(\rho_3-1)/4}\right] \pmod{2^{\lambda+3}}. \end{aligned}$$

But the quantity in brackets is divisible by 8. In fact if both k and ρ_3 are of the form $4x-1$ it becomes

$$(3\rho_3 + 1)(k - 1) + 4 \equiv 0 \pmod 8.$$

In the opposite case we have

$$3(k-1)(\rho_3 - 1) \equiv 0 \pmod 8.$$

Hence in both cases D_2 is divisible by $2^{\lambda+3}$. This completes the proof of the theorem.

THEOREM 3. *Let k be an odd integer, then*

$$A_k(n) = A_{2k}(4n + (k^2 - 1)/8). \tag{2.17}$$

Proof. Since

$$A_{2k}(n_1) = \sum_{(\rho_1)} \exp\left[g_{n_1}(\rho_1, 2k)\frac{\pi i}{24k}\right], \tag{2.18}$$

while

$$A_k(n) = \sum_{(\rho_2)} \exp\left[2f_n(\rho_2, k)\frac{\pi i}{24k}\right], \tag{2.19}$$

where for brevity we take n_1 for $4n+(k^2-1)/8$, it suffices to show that a correspondence may be set up between ρ_1 and ρ_2 so that corresponding terms of (2.18) and (2.19) are equal. We show that the correspondence is simply

$$\rho_1 \equiv \rho_2/2 \pmod{k}. \tag{2.20}$$

To this effect we define D_3 by

$$D_3 = 2f_n(\rho_2, k) - g_{n_1}(\rho_1, 2k) \tag{2.21}$$

and show that $D_3 \equiv 0 \pmod{48k}$.

We first show that D_3 is divisible by $3k$. In case 3 is prime to k, the fact that $D_3 \equiv 0 \pmod{k}$ is an immediate consequence of (1.12) and (1.15). In case 3 divides k we use (1.11) and (1.14). In either case

$$D_3 \equiv 2\nu\rho_2 + 2\bar{\rho}_2 - \nu_1\rho_1 - \bar{\rho}_1 \pmod{k \text{ or } 3k}.$$

But

$$\nu_1 = 1 - 24n_1 = 1 - 96n - 3(k^2 - 1) \equiv 4(1 - 24n) \equiv 4\nu \pmod{3k}. \tag{2.22}$$

So that

$$D_3 \equiv (\rho_2 - 2\rho_1)\left(2\nu - \frac{1}{\rho_1\rho_2}\right) \pmod{k \text{ or } 3k}.$$

By (2.20) the first factor is divisible by k while in case $3|k$, the second factor contains 3 by (2.20), since $\nu \equiv 1 \pmod{3}$. Hence $D_3 \equiv 0 \pmod{3k}$.

We must now show that D_3 is divisible by 16. Using (1.13) and (1.16) in (2.21) we find

$$D_3 \equiv 4k\left(\frac{-\rho_2}{k}\right) + 2k - 6 - \nu_1\rho_1 - 4k\left(\frac{-2k}{\rho_1}\right) - 4\rho_1 - 6k\rho_1 - \bar{\rho}_1 \pmod{16}.$$

In view of (2.20) and (2.22) we may substitute for ρ_2 and ν_1 and obtain

$$D_3 \equiv 4\left\{k\left[\left(\frac{-2\rho_1}{k}\right) - \left(\frac{-2k}{\rho_1}\right)\right] + 2\,\frac{(k-3)(k\rho_1+1)}{4} + 2\bar{\rho}_1\,\frac{k^2\rho_1^2 - 1}{8} + 2\right\} \pmod{16}.$$

Considering the quantity inside the braces modulo 4 and replacing all odd factors of even numbers by unity we obtain

$$D_3 \equiv 4\left\{\left(\frac{-2\rho_1}{k}\right) - \left(\frac{-2k}{\rho_1}\right) + 1 - (-1)^{(\rho_1+1)(k+1)/4} + 1 - \left(\frac{2}{k\rho_1}\right) + 2\right\}$$

$$\equiv 4\left\{\left(\frac{-2}{k}\right)-\left(\frac{-2}{\rho_1}\right)(-1)^{(\rho_1-1)(k-1)/4} + (-1)^{(\rho_1-1)(k-1)/4}\left(\frac{-1}{k\rho_1}\right)-\left(\frac{2}{k\rho_1}\right)\right\}$$

$$\equiv 4\left\{\left(\frac{-2}{k}\right)-\left(\frac{-2}{\rho_1}\right)+\left(\frac{-1}{k\rho_1}\right)-\left(\frac{2}{k\rho_1}\right)\right\}$$

$$\equiv 4\left\{\left(\frac{-1}{k}\right)-\left(\frac{2}{\rho_1}\right)\right\}\left\{\left(\frac{-1}{\rho_1}\right)+\left(\frac{2}{k}\right)\right\} \equiv 0 \pmod{16}.$$

This proves Theorem 3.

Another form of Theorem 3, which exhibits it as a multiplication theorem, is

THEOREM 4. *If k is odd, then*

$$A_2(n_1)A_k(n_2) = A_{2k}(n_3) \tag{2.23}$$

where

$$n_3 \equiv 4n_2 + kn_1 + (k^2-1)/8 \pmod{2k}.$$

This follows easily from the following

LEMMA 1. *If k and n are arbitrary integers*

$$A_{2k}(n) = -A_{2k}(n+k).$$

Proof. This is an immediate consequence of applying the identity

$$e^{-2n\rho\pi i/2k} = -e^{-2(n+k)\rho\pi i/2k}$$

where ρ is odd, to the definition (1.2) of $A_k(n)$.

To prove Theorem 4 we need only to observe that if we apply Lemma 1 n_1 times with $n=4n_2+(k^2-1)/8$, we obtain from Theorem 3

$$A_{2k}\left(4n_2+kn_1+\frac{k^2-1}{8}\right) = (-1)^{n_1}A_k(n_2) = A_2(n_1)A_k(n_2).$$

Another form of Theorem 3 states that if k is odd,

$$A_{2k}(n) = (-1)^n\left(\frac{2}{k}\right)A_k\left(\frac{8n+1}{32}\right).$$

This follows readily from Theorem 4.

Theorems 1, 2, and 4 may be used to express $A_k(n)$ in terms of A's whose subscripts are powers of primes dividing k. This is illustrated in the following examples.

Example I. Express $A_{35}(23)$ in terms of A_5 and A_7. By (2.1), $n_3=23\equiv 25n_2+49n_1-73/24 \pmod{35}$. This gives two congruences

$$49n_1 \equiv 0 \pmod 5$$

and

$$25n_2 \equiv 3 \pmod 7.$$

Hence $n_1\equiv 0 \pmod 5$ and $n_2\equiv 6 \pmod 7$. Therefore

$$A_{35}(23) = A_5(0)A_7(6).$$

Example II. Express $A_{30}(17)$ in terms of A_2, A_3, and A_5. By Theorem 4 with $k=15$ we have

$$n_3 = 17 \equiv 4n_2 + 15n_1 + 28 \pmod{30},$$

whence

$$n_2 \equiv 1 \pmod{15},$$

$$n_1 \equiv 1 \pmod 2,$$

so that

$$A_{30}(17) = A_2(1)A_{15}(1).$$

Applying Theorem 1 to $A_{15}(1)$ as in Example I, we find

$$A_{15}(1) = A_3(2)A_5(2),$$

whence

$$A_{30}(17) = A_2(1)A_3(2)A_5(2).$$

Example III. Express $A_{36}(13)$ in terms of A_4 and A_9. By Theorem 2, $A_{36}(13)=-A_9(n_1)A_4(n_2)$, where $13\equiv 81n_2+16n_1-4$, so that

$$n_1 \equiv 5 \pmod 9,$$

$$n_2 \equiv 1 \pmod 4$$

and we have

$$A_{36}(13) = - A_4(1)A_9(5).$$

3. **The evaluation of** $A_q(n)$. From the results of the preceding section it is clear that questions concerning the actual value of $A_k(n)$ or merely the order of magnitude of $A_k(n)$ may be reduced to the corresponding questions about $A_q(n)$, where q is a power of a prime. Three cases present themselves quite naturally, namely those in which (I) $q=p^\alpha$, where p is a prime >3 and $\alpha\geqq 1$, (II) q is a power of 3, (III) q is a power of 2. In all cases the number $A_q(n)$ may be expressed in terms of generalized Kloosterman sums of the type

$$\sum_{(s)} \chi(s) e^{2\pi i(as+\bar{s})/q},$$

where χ is a quadratic character, and $s\bar{s}\equiv 1 \pmod{q}$.

The problem of evaluating these sums has been solved in case q is a prime by Salié. In case q is a power of an odd prime the sums may be easily evaluated if use is made of Salié's discussion (§§2, 3) of the original Kloosterman sum. In fact the introduction of the character has no influence on the argument until the very last stage where it results in changing some of the cosines to sines or vice versa. We give therefore without further comment the following lemmas.

LEMMA* 2. *Let $q=p^\alpha$, where p is an odd prime. Then if s runs over the numbers less than and prime to q, the sum*

$$\sum_{(s)} \left(\frac{s}{q}\right) e^{2\pi i(as+\bar{s})/q} = \begin{cases} 0 \text{ if } a \text{ is a non-residue of } q \text{ prime to } q, \\ i^{((q-1)/2)^2} 2q^{1/2} \cos \dfrac{4\pi\theta}{q} \text{ if } a \equiv \theta^2 \pmod{q}, \ a \text{ prime to } q, \\ 0 \text{ if } a \text{ is divisible by } p \text{ and } \alpha > 1, \\ i^{((q-1)/2)^2} q^{1/2} \text{ if } a \text{ is divisible by } p \text{ and } \alpha = 1. \end{cases}$$

LEMMA 3. *If $q=3^\alpha$, $\alpha>1$ and if $a\equiv 1 \pmod{3}$, then*

$$\sum_{(s)} \left(\frac{s}{3^{\alpha-1}}\right) e^{2\pi i(as+\bar{s})/q} = i^{((q-1)/2)^2+1} \left(\frac{\theta}{3}\right) 2q^{1/2} \sin \frac{4\pi\theta}{q},$$

where $\theta^2\equiv a \pmod{q}$.

To apply these lemmas to the evaluation of $A_q(n)$ where q is odd we separate two cases.

Case I. $q=p^\alpha$, where p is a prime >3. Returning to the definition (1.8) of $f_n(\rho)$, we have from congruences (1.12) and (1.13)

$$f_n(\rho) \equiv 0 \pmod{3}$$

and

$$f_n(\rho) \equiv q - 3 + 2q\left(\frac{-\rho}{q}\right) \pmod{8}.$$

If $q\equiv 1 \pmod{4}$, $q-3\equiv -2q(2/q) \pmod{8}$. Hence $f_n(\rho)$ is an even or odd multiple of 12 according as $(2\rho/q)=+1$, or -1. If $(2\rho/q)=+1$ we may write

$$f_n(\rho) \equiv 24t(\rho) \pmod{24q},$$

while, if $(2\rho/q)=-1$,

* Cf. Salié, loc. cit., equations (54) and (57), p. 102, in case q is a prime. For q a power of a prime compare (32) and (33), p. 97.

$$f_n(\rho) \equiv 24t(\rho) + 12q \pmod{24q},$$

where, in both cases,

$$(3.1) \qquad t \equiv t(\rho) \equiv f_n(\rho)/24 \equiv (\nu\rho + \bar{\rho})/24 \pmod{q}$$

by (1.11). Hence for a given ρ we get a term of (1.6) of the form $e^{2\pi it/q}$ or $e^{2\pi it/q}e^{\pi i}$, so that if $q \equiv 1 \pmod 4$,

$$A_q(n) = \sum_{(\rho)} \left(\frac{2\rho}{q}\right) e^{2\pi it/q},$$

where t is defined by (3.1).

If $q \equiv -1 \pmod 4$, $f_n(\rho) \equiv 0 \pmod 3$ and

$$f_n(\rho) \equiv -6q\left(\frac{2\rho}{q}\right) \pmod 8.$$

Hence, in this case,

$$f_n(\rho) = 24t - 6q\left(\frac{2\rho}{q}\right) \pmod{24q},$$

where t is given by (3.1) and the typical term of (1.6) is, in this case,

$$e^{2\pi it/q} \exp\left[-\left(\frac{2\rho}{q}\right)\frac{\pi i}{2}\right] = -i\left(\frac{2\rho}{q}\right)e^{2\pi it/q}.$$

So that whether $q \equiv +1$ or $-1 \pmod 4$, we have

$$(3.2) \quad A_q(n) = (-i)^{((q-1)/2)^2} \sum_{(\rho)} \left(\frac{2\rho}{q}\right) e^{2\pi it/q}, \text{ where } t \equiv (\nu\rho + \bar{\rho})/24 \pmod{q}.$$

In order to apply Lemma 2 to (3.2) we set $24\rho \equiv s \pmod q$, so that

$$t \equiv \frac{\nu}{24^2} s + \bar{s} \pmod{q}$$

and

$$\left(\frac{2\rho}{q}\right) = \left(\frac{3s}{q}\right) = \left(\frac{3}{q}\right)\left(\frac{s}{q}\right).$$

With this change in notation (3.2) becomes

$$A_q(n) = (-i)^{((q-1)/2)^2} \left(\frac{3}{q}\right) \sum_{(s)} \left(\frac{s}{q}\right) e^{2\pi i(as+\bar{s})/q}, \text{ where } a \equiv \nu/24^2 \pmod{q}.$$

Applying Lemma 2, we obtain

THEOREM 5. *If* $q=p^\alpha$, p *a prime* >3, $\alpha \geqq 1$ *and* $\nu = 1-24n$, *then*

$$A_q(n) = \begin{cases} 0 \text{ if } \nu \text{ is non-residue* of } q, \text{ prime to } q, \\ 2\left(\dfrac{3}{q}\right) q^{1/2} \cos \dfrac{4\pi m}{q} \text{ if } \nu \equiv (24m)^2 \pmod{q}, \text{ prime to } q, \\ 0 \text{ if } \nu \equiv 0 \pmod{p} \text{ and } \alpha > 1, \\ \left(\dfrac{3}{q}\right) q^{1/2} \text{ if } \nu \equiv 0 \pmod{p} \text{ and } \alpha = 1. \end{cases}$$

Case II. $q=3^\beta$, $\beta \geqq 1$. First let β be even. Then from (1.11) and (1.13) we have

$$f_n(\rho) \equiv 8t(\rho) \pmod{3^{\beta+1}}.$$

$$t = t(\rho) \equiv (\nu\rho + \bar{\rho})/8 \pmod{3^{\beta+1}}, \tag{3.3}$$

so that $\exp[f_n(\rho)\pi i/12q] = e^{2\pi i t/3^{\beta+1}}$.

If β is odd, then

$$f_n(\rho) \equiv 8t - 6q\left(\frac{2\rho}{q}\right) \pmod{8\cdot 3^{\beta+1}}.$$

Hence in this case

$$\exp\,[f_n(\rho)\pi i/12q] = -\,i\left(\frac{2\rho}{q}\right)e^{2\pi i t/3^{\beta+1}},$$

where t is given again by (3.3). Hence in both cases we have

$$A_q(n) = (-\,i)^{((q-1)/2)^2}\sum_{(\rho)}\left(\frac{2\rho}{q}\right)e^{2\pi i t(\rho)/3^{\beta+1}}, \tag{3.4}$$

where ρ runs over the numbers prime to 3 and less than 3^β. Since $t(\rho+3^\beta)\equiv t(\rho)$ (mod $3^{\beta+1}$), if $\bar{\rho}$ is replaced by $\bar{\rho}-3^\beta\bar{\rho}^2$ we may let ρ run up to $3^{\beta+1}$ in (3.4) and obtain $3A_q(n)$. At the same time we replace ρ by $s/8$ (mod $3^{\beta+1}$), so that

$$t(\rho) \equiv \frac{\nu}{8^2}\,s + \bar{s} \pmod{3^{\beta+1}}$$

and $(2\rho/q)=(s/q)$. Hence (3.4) becomes

$$A_q(n) = \tfrac{1}{3}(-\,i)^{((q-1)/2)^2}\sum_{(s)}\left(\frac{s}{q}\right)e^{2\pi i(as+\bar{s})/3^{\beta+1}},$$

* This condition should not be confused with $(\nu/q)=-1$. We mean that no solution exists of the congruence $x^2\equiv\nu$ (mod q).

where $a \equiv \nu/8^2 \pmod{3^{\beta+1}}$. Since $\nu \equiv 1 \pmod 3$, we may apply Lemma 3 with $\alpha = \beta+1$ and obtain the following:

THEOREM 6. *Let* $\nu = 1 - 24n$, *then*

$$A_{3^\beta}(n) = (-1)^{\beta+1}\left(\frac{m}{3}\right)\frac{2}{3^{1/2}}(3^\beta)^{1/2}\sin\frac{4\pi m}{3^{\beta+1}},$$

where $(8m)^2 \equiv \nu \pmod{3^{\beta+1}}$.

Case III. $q = 2^\lambda$. In this case we shall evaluate $A_q(n)$ directly without passing to a generalized Kloosterman sum, since the introduction of the appropriate quadratic character into Salié's discussion of the corresponding Kloosterman sum cannot be accomplished without considerable reconstruction. The method of proof is similar to that used by Salié and Estermann.

THEOREM 7. *If* $\lambda \geqq 0$,

$$A_{2^\lambda}(n) = (-1)^\lambda\left(\frac{-1}{m}\right)(2^\lambda)^{1/2}\sin\frac{4\pi m}{2^{\lambda+3}},$$

where m is an integer $\equiv \nu^{1/2}/3 \pmod{2^{\lambda+3}}$.

Proof. For brevity define u and v by

$$u = \left[\frac{\lambda}{2}\right] + 2, \qquad v = \left[\frac{\lambda+1}{2}\right] - 2$$

so that $\lambda = u+v$. The numbers ρ which are less than 2^λ and odd may be represented by

$$\rho = \tau + 2^u h \qquad \begin{cases} \tau = 1, 3, 5, \cdots, 2^u - 1, \\ h = 0, 1, 2, \cdots, 2^v - 1. \end{cases}$$

Hence $A_{2^\lambda}(n)$ may be written as a double sum

$$(3.5) \qquad A_{2^\lambda}(n) = \sum_{(\tau)}\sum_{h=0}^{2v-1}\exp\left[g_n(\tau + 2^u h)\pi i/12\cdot 2^\lambda\right].$$

For $h_1 \neq h_2$ we consider the difference

$$\Delta_\tau = g_n(\tau + 2^u h_1) - g_n(\tau + 2^u h_2).$$

By (1.15), $\Delta_\tau \equiv 0 \pmod 3$. Assuming that $\lambda > 4$ so that $u > 3$, and $2u \geqq \lambda + 3$ we find from (1.16) that

$$(3.6) \qquad \Delta_\tau \equiv 2^u(h_1 - h_2)\left\{\nu - \frac{1}{\tau^2}\right\} \pmod{2^{\lambda+3}}.$$

Since

$$(3.7) \qquad \nu \equiv \bar{\nu} \equiv 1 \pmod 8,$$

it follows that ν is a quadratic residue of $2^{\lambda+3}$. In view of (3.6), $\Delta_\tau \equiv 0 \pmod{2^{u+3}}$ for all τ. We proceed to arrange the values of τ into sets according to the highest power of 2 dividing $\tau^2 - \bar{\nu}$.

If $\tau^2 \not\equiv \bar{\nu} \pmod{2^4}$, τ will be said to belong to set 1. For such a τ, $\Delta_\tau \equiv 0 \pmod{2^{u+3}}$ for all pairs (h_1, h_2), but since each $h < 2^v$, $h_1 - h_2$ is never divisible by 2^v, that is, $\Delta_\tau \not\equiv 0 \pmod{2^{\lambda+3}}$. This means that those terms of (3.5) for which τ belongs to set 1 correspond to g's of the form

$$g_n(\tau + 2^u h) = c_\tau + 3M_{\tau,h}2^{u+3} \pmod{24 \cdot 2^\lambda},$$

where $M_{\tau,h}$ runs with h over the numbers $0, 1, 2, \cdots, 2^v - 1$ in some order. Hence the contribution to (3.5) from any member τ of set 1 is

$$\exp\left[\frac{c_\tau \pi i}{12 \cdot 2^\lambda}\right]\sum_{h=0}^{2^v-1} \exp\frac{2\pi i h}{2^v} = 0.$$

Hence we need only consider those τ's which do not belong to set 1. For such numbers $\tau^2 \equiv \bar{\nu} \pmod{2^4}$. For any $h_1 < 2^{v-1}$ there exists an $h_2 = h_1 + 2^{v-1}$ such that $(h_1 - h_2) \equiv 0 \pmod{2^{v-1}}$ and the corresponding difference Δ_τ is divisible by $3 \cdot 2^u \cdot 2^{v-1} \cdot 2^4 = 24 \cdot 2^\lambda$. Since the corresponding terms of (3.5) make precisely the same contribution to $A_{2^\lambda}(n)$ we may contract (3.5) to read

$$(3.8) \qquad A_{2^\lambda}(n) = 2\sum_{(\tau)} \sum_{h=0}^{2^{v-1}-1} \exp\left[g_n(\tau + 2^u h)\pi i/12 \cdot 2^\lambda\right],$$

where now the outer sum extends over the values of τ for which

$$\tau^2 \equiv \bar{\nu} \pmod{2^4}.$$

If $\tau^2 \not\equiv \bar{\nu} \pmod{2^5}$, τ will be said to belong to set 2. For such a τ, $\Delta_\tau \equiv 0 \pmod{2^{u+4}}$, but $\Delta_\tau \not\equiv 0 \pmod{2^{\lambda+3}}$, since now $h_1 - h_2$ is never divisible by 2^{v-1}. This means that the terms of (3.8) belonging to a fixed number τ of set 2 contribute nothing to $A_{2^\lambda}(n)$. We may therefore ignore all τ's but those for which $\tau^2 \equiv \bar{\nu} \pmod{2^5}$. Moreover if for any such τ, $h_2 = h_1 + 2^{v-2}$ the corresponding Δ_τ will be divisible by $3 \cdot 2^u \cdot 2^{v-2} \cdot 2^5 = 24 \cdot 2^\lambda$ so that the corresponding contributions to (3.8) are identical. Hence

$$A_{2^\lambda}(n) = 2^2\sum_{(\tau)} \sum_{h=0}^{2^{v-2}-1} \exp\left[g_n(\tau + 2^u h)\pi i/12 \cdot 2^\lambda\right],$$

where now $\tau^2 \equiv \bar{\nu} \pmod{2^5}$. Repeating the argument we may reduce the terms of the inner sum to a single term corresponding to $h = 0$ and obtain

$$(3.9) \qquad A_{2^\lambda}(n) = 2^v \sum_{(\tau)} \exp\left[g_n(\tau)\pi i/12 \cdot 2^\lambda\right],$$

where

$$(3.10) \qquad \tau^2 \equiv \bar{\nu} \pmod{2^{v+3}}.$$

At this point we separate two cases according to the parity of λ. If λ is odd, $u=v+3$, and since τ is chosen from the odd numbers $<2^u=2^{v+3}$ the congruence (3.10) has four solutions

$$\tau \equiv \pm \gamma, \quad \pm (\gamma + 2^{v+2}) \quad \pmod{2^{v+3}}$$

where $\gamma^2 \equiv \bar{\nu} \pmod{2^{\lambda+3}}$. If we consider the difference

$$D_1 = g_n(\gamma + 2^{v+2}) - g_n(\gamma),$$

we find by (1.15) that $D_1 \equiv 0 \pmod 3$, while by (1.16)

$$\begin{aligned} D_1 &\equiv 2^{v+2}\left\{\nu - \frac{1}{\gamma(\gamma + 2^{v+2})}\right\} \\ &\equiv 2^{v+2}\{\nu\gamma^2 - 1 + \nu\gamma 2^{v+2}\}(\gamma^2 + 2^{v+2}\gamma)^{-1} \pmod{2^{\lambda+3}}. \end{aligned}$$

Since $\nu\gamma^2 - 1 \equiv 0 \pmod{2^{v+2}}$ the last factor may be taken modulo $2^{\lambda+3-2v-4}=4$, and is $\equiv 1 \pmod 4$ since $v \geqq 0$. Hence

$$D_1 \equiv 2^{\lambda+1}\bar{\gamma} \pmod{2^{\lambda+3}},$$

and

$$g_n(\gamma + 2^{v+2}) \equiv g_n(\gamma) + 3c2^{\lambda+1} \pmod{24\cdot 2^\lambda}$$

where c is an integer $\equiv 1/3\gamma \equiv -\gamma \pmod 4$. Since $g(\rho) = -g(-\rho) \pmod{24\cdot 2^\lambda}$, (3.9) becomes

$$\begin{aligned} A_2{}^\lambda(n) &= 2^{v+1}\{\cos[g_n(\gamma)\pi/12\cdot 2^\lambda] + \cos[g_n(\gamma)\pi/12\cdot 2^\lambda - \gamma\pi/2]\} \\ &= 2^{v+2}\cos\frac{\pi\gamma}{4}\cos[g_n(\gamma)\pi/12\cdot 2^\lambda - \gamma\pi/4]. \end{aligned}$$

But $v+2=(\lambda+1)/2$ and $\cos \pi\gamma/4 = (2/\gamma)(1/2^{1/2})$, so that we have

$$(3.11) \qquad A_2{}^\lambda(n) = \left(\frac{2}{\gamma}\right)(2^{\lambda/2})\cos[g_n(\gamma)\pi/12\cdot 2^\lambda - \pi\gamma/4].$$

Before proceeding further we take up the case of λ even. In this case $u=v+4$. Since $\tau<2^u=2^{v+4}$ there are 8 values of τ which are needed in (3.9). These are

$$\tau \equiv \pm (\gamma + j2^{v+2}) \pmod{2^{v+3}} \qquad (j = 0, 1, 2, \text{ and } 3).$$

Now if

$$D_2 = g_n(\gamma + j2^{v+2}) - g_n(\gamma),$$

then $D_2 \equiv 0 \pmod 3$ and

$$D_2 \equiv j2^{v+2}\left\{\nu - \frac{1}{\gamma(\gamma + j2^{v+2})}\right\}$$
$$\equiv \bar{\gamma}j^2 2^{2v+4} \equiv 9\gamma j^2 2^\lambda \pmod{2^{\lambda+3}}.$$

Therefore

$$g_n(\gamma + j2^{v+2}) \equiv g_n(\gamma) + 9\gamma j^2 2^\lambda \pmod{3\cdot 2^{\lambda+3}},$$

so that if

$$F(j) = \exp\,[g_n(\gamma + j2^{v+2})\pi i/12\cdot 2^\lambda] = \exp\,[g_n(\gamma)\pi i/12\cdot 2^\lambda]e^{3\gamma j^2\pi i/4},$$

then $F(0) = -F(2)$ and $F(1) = F(3)$, and (3.9) becomes

$$A_{2^\lambda}(n) = 2^{v+2}\cos\,[g_n(\gamma)\pi/12\cdot 2^\lambda + 3\gamma\pi/4]$$
$$= -\,(2^{\lambda/2})\cos\,[g_n(\gamma)\pi/12\cdot 2^\lambda - \gamma\pi/4].$$

In view of (3.11) we may combine the results for even and odd by writing

$$(3.12)\qquad A_{2^\lambda}(n) = (-1)^{\lambda+1}\left(\frac{2^\lambda}{\gamma}\right)(2^{\lambda/2})\cos\,[g_n(\gamma)\pi/12\cdot 2^\lambda - \gamma\pi/4].$$

To evaluate this cosine we refer once more to (1.15) and (1.16) and write $g_n(\gamma) \equiv 0 \pmod 3$ and

$$g_n(\gamma) \equiv \nu\gamma + \bar{\gamma} - 6\cdot 2^\lambda\left(\frac{-2^\lambda}{\gamma}\right) + 3\gamma 2^\lambda \pmod{2^{\lambda+3}},$$

so that if we define the integer m by

$$m \equiv (\nu\gamma + \bar{\gamma})/6 \equiv \nu^{1/2}/3 \pmod{2^{\lambda+3}},$$

we have

$$\cos\,[g_n(\gamma)\pi/12\cdot 2^\lambda - \gamma\pi/4] = \cos\left\{\frac{4\pi m}{2^{\lambda+3}} - \left(\frac{-2^\lambda}{\gamma}\right)\frac{\pi}{2}\right\}$$
$$= \left(\frac{-2^\lambda}{\gamma}\right)\sin\frac{4\pi m}{2^{\lambda+3}}\cdot$$

Hence substituting into (3.12) we have

$$A_{2^\lambda}(n) = (-1)^{\lambda+1}\left(\frac{-1}{\gamma}\right)(2^{\lambda/2})\sin\frac{4\pi m}{2^{\lambda+3}},$$

where m is an integer $\equiv \nu^{1/2}/3 \pmod{2^{\lambda+3}}$. But

$$\left(\frac{-1}{\gamma}\right) = \left(\frac{-1}{3}\right)\left(\frac{-1}{m}\right) = -\left(\frac{-1}{m}\right),$$

so the theorem follows, when $\lambda > 4$.

As a matter of fact the theorem is true when $\lambda \leqq 4$. This may be verified by merely consulting the tables of $A_k(n)$.

The following corollary, which is a consequence of Theorems 1, 2, 3, and 5, is especially useful in applying series (1.1). The proof, which involves the separation of three cases, is left to the reader.

COROLLARY. *If k is divisible by the prime p and if $A_p(n) = 0$, then $A_k(n) = 0$.*

4. **Estimates and remainders.** In this section we apply the preceding results to the discussion of the order of magnitude of $A_k(n)$ and to the estimation of the errors committed by taking only the first N terms of the Hardy-Ramanujan and Rademacher series.

THEOREM 8. *Let ω be the number of distinct odd prime factors of k. Then*

$$|A_k(n)| < 2^\omega k^{1/2}.$$

Proof. Let

$$k = 2^\lambda p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_\omega^{\alpha_\omega}$$

be the decomposition of k into its prime factors. By §2, there exists a set $(n_0, n_1, n_2, \cdots, n_\omega)$ such that

$$A_k(n) = A_{2^\lambda}(n_0) A_{p_1^{\alpha_1}}(n_1) \cdots A_{p_\omega^{\alpha_\omega}}(n_\omega). \tag{4.1}$$

By theorem 7

$$|A_{2^\lambda}(n_0)| < 2^{\lambda/2},$$

while by Theorems 5 and 6

$$|A_{p^\alpha}(n)| < 2p^{\alpha/2}.$$

Hence the theorem follows from (4.1).

Since, for every $\epsilon > 0$

$$2^{\omega(k)} = O(k^\epsilon),$$

we have as an immediate consequence

$$A_k(n) = O(k^{1/2+\epsilon}). \tag{4.2}$$

Using this result it is possible to prove the following:

THEOREM 9. *Let w be any number greater than $2\pi(2/3)^{1/2} = 5.130 \cdots$, then for all sufficiently large n, $p(n)$ is the nearest integer to the first $[wn^{1/2}/\log n]$ terms of the Hardy-Ramanujan series.*

This theorem is similar to a result of Hardy and Ramanujan* based on

* Loc. cit., pp. 107–108, §6.3.

$A_k(n) = O(k)$ in which w is replaced by $4\pi(2/3)^{1/2}$, and may be proved in the same way. However it is possible to prove somewhat more than (4.2).

THEOREM 10. *For every $\epsilon > 0$ there exists a K such that for all n and for all $k > K$*

$$|A_k(n)| < k^{1/2+(1+\epsilon)\log 2/\log\log k}.$$

Proof. This follows easily from a theorem of Wigert* to the effect that for every ϵ there exists a K such that if $k > K$

$$\tau(k) < 2^{(1+\epsilon)\log k/\log\log k},$$

where $\tau(k)$ is the number of divisors of k, and from the trivial inequality

$$2^{\omega(k)} \leqq \tau(k).$$

In contrast to this theorem we prove:

THEOREM 11. *For every $\epsilon > 0$ there exist infinitely many values of n and k for which*

$$|A_k(n)| > .11367k^{1/2+(1-\epsilon)\log 2/\log\log k}.$$

Proof. Let p be a prime > 3 and let n_p be an integer $\equiv -35/24 \pmod p$. Then

$$\nu_p = 1 - 24n_p \equiv 6^2 \pmod p.$$

Applying Theorem 5 with $q = p$ and $n = n_p$ we find $m \equiv (p^2 - 1)/4 \pmod p$ and

$$(4.3) \qquad |A_p(n_p)| = 2p^{1/2}\left|\cos\frac{(p^2-1)\pi}{p}\right| = 2p^{1/2}\cos\frac{\pi}{p} > 2p^{1/2}\left(1 - \frac{\pi^2}{2p^2}\right).$$

Let p_j denote the jth prime $\geqq 2$. By Theorem 1 there exists an n such that

$$A_k(n) = A_2(0)A_3(0)A_5(n_5)\cdots A_{p_t}(n_{p_t}),$$

where $k = 2\cdot 3\cdot 5\cdot 7\cdot\,\cdots\,\cdot p_t$ and where t will be determined later. Applying (4.3) we find

$$|A_k(n)| > \frac{A_3(0)}{6^{1/2}}\, k^{1/2}2^{t-2}\prod_{j=3}^{t}\left(1 - \frac{\pi^2}{2p_j^2}\right)$$

$$> \frac{A_3(0)}{6^{1/2}}\, k^{1/2}2^{t-2}\,\frac{\prod_{j=1}^{\infty}(1 - \pi^2/2(2j-1)^2)}{(1-\pi^2/2)(1-\pi^2/18)(1-\pi^2/162)}$$

$$> .12044k^{1/2}\cdot 2^t\left|\cos(\pi^2/8^{1/2})\right|$$

$$(4.4) \qquad |A_k(n)| > .11367k^{1/2}\cdot 2^t.$$

* Archiv för Matematik, Astronomi, och Fysik, vol. 3 (1906–1907), no. 18. See also Landau Handbuch, vol. 1, p. 220.

As to the factor 2^t, we have by the prime number theorem

$$\log 2^t = t\cdot\log 2 = \pi(p_t)\log 2 \sim p_t \log 2/\log p_t.$$

But

$$p_t \sim \sum_{j=1}^{t} \log p_t = \log k.$$

Hence

$$\log 2^t \sim \frac{\log 2 \log k}{\log\log k}.$$

Therefore if $\epsilon > 0$ is given

$$\log 2^t > \frac{(1-\epsilon)\log 2\log k}{\log\log k}$$

for all sufficiently large values of t and k. In other words there are infinitely many k's for which

$$2^t > k^{(1-\epsilon)\log 2/\log\log k}.$$

From this and (4.4) the theorem follows at once.

In the subsequent discussion however we need an estimate of $A_k(n)$ for all values of k and n. The one we shall use is given by the following:

THEOREM 12. *For all n and k*

$$|A_k(n)| < 2k^{5/6}. \tag{4.5}$$

Proof. Let $\omega(k)$ denote as before the number of odd prime factors of k and let P_j be the product of the first j odd primes. The number k being given, there exists a j such that

$$P_j \leqq k < P_{j+1}.$$

This means that

$$\omega(k) \leqq j. \tag{4.6}$$

Suppose for the moment that $k \geqq 105 = P_3$. Then since $P_{j+1} > k$, we have $j \geqq 3$. Hence

$$k \geqq P_j \geqq 3\cdot 5\cdot 7\cdot 11^{j-3}$$

or

$$j \leqq .96026 \log_{10} k + 1.05914.$$

Therefore in view of (4.6)

$$2^{\omega(k)} \leqq 2^j < 2.0836\, k^{.28907} < 2k^{1/3}$$

since $k>3$. It follows by Theorem 8 that

$$|A_k(n)| < 2k^{5/6} \text{ for } k \geqq 105.$$

If $15 \leqq k < 105$, then $2^{\omega(k)} \leqq 4 < 2\cdot(15)^{1/3}$ and $|A_k(n)| < 2(15)^{1/3}\cdot k^{1/2} \leqq 2k^{5/6}$. For $1 \leqq k < 15$, $2^{\omega(k)} \leqq 2$ so that $|A_k(n)| < 2\ k^{1/2} < 2k^{5/6}$. Hence the inequality (4.5) holds for all k.

It is clear that an infinite number of inequalities similar to (4.5) but with smaller powers ($>1/2$) of k may be established in the same way (for example $|A_k(n)| < 3k^{3/4}$), but only at the expense of larger constant coefficients.

We now consider the remainder of Rademacher's convergent series for $p(n)$.

$$p(n) = \frac{1}{\pi 2^{1/2}} \sum_{k=1}^{\infty} A_k(n) k^{1/2} \frac{d}{dn}\left(\frac{\sinh\ (c\lambda_n/k)}{\lambda_n}\right), \tag{4.7}$$

where

$$c = \pi(2/3)^{1/2}, \qquad \lambda_n = (n - 1/24)^{1/2}.$$

Introducing the notation

$$c\lambda_n = \mu_n = \mu = \frac{\pi}{6}(24n-1)^{1/2}, \tag{4.8}$$

$$A_k(n) = A_k^*(n) k^{1/2}, \tag{4.9}$$

we may write (4.7) in the form

$$p(n) = \frac{12^{1/2}}{24n-1} \sum_{k=1}^{N} A_k^*(n) \left\{\left(1 - \frac{k}{\mu}\right) e^{\mu/k} + \left(1 + \frac{k}{\mu}\right) e^{-\mu/k}\right\} + R_N(n), \tag{4.10}$$

where the remainder $R_N(n)$ may be written after expanding the exponentials and collecting

$$R_N(n) = \frac{4(12)^{1/2}}{24n-1} \sum_{k=N+1}^{\infty} A_k^*(n) \sum_{j=1}^{\infty} \frac{j(\mu/k)^{2j}}{(2j+1)!}. \tag{4.11}$$

By Theorem 12

$$|A_k^*(n)| < 2k^{1/3}. \tag{4.12}$$

Therefore if we eliminate n from (4.11) by means of (4.8) we obtain

$$\begin{aligned} |R_N(n)| &< \frac{4\pi^2 3^{1/2}}{9\mu^2} \sum_{k=N+1}^{\infty} k^{1/3} \sum_{j=1}^{\infty} \frac{j(\mu/k)^{2j}}{(2j+1)!} \\ &< \frac{4\pi^2 3^{1/2}}{9\mu^2} \int_N^{\infty} \sum_{j=1}^{\infty} \frac{j(\mu/x)^{2j} x^{1/3}}{(2j+1)!}\, dx. \end{aligned}$$

Setting $\mu/x=t$ and defining r by

$$r=\mu/N,$$

we obtain on eliminating x

$$|R_N(n)| < \frac{4\pi^2 3^{1/2}}{9\mu^{2/3}}\int_0^r \sum_{j=1}^{\infty}\frac{jt^{2j-7/3}}{(2j+1)!}dt$$

$$(4.13) \qquad < \frac{N^{-2/3}\pi^2}{3^{1/2}}\sum_{j=0}^{\infty}\frac{r^{2j}}{(2j+3)(3j+1)(2j+1)!}$$

$$< \frac{N^{-2/3}\pi^2}{3^{1/2}}\left\{\frac{1}{3}+\sum_{j=1}^{\infty}\frac{r^{2j}}{(2j+3)!}\right\}.$$

So that

$$(4.14) \qquad |R_N(n)| < \frac{N^{-2/3}\pi^2}{3^{1/2}}\left\{\frac{\sinh r}{r^3}+\frac{1}{6}-\frac{1}{r^2}\right\} = N^{-2/3}F(r).$$

We give a few values of $F(r)$ for typical values of r.

r	$F(r)$	r	$F(r)$
1.0	1.9480	3.5	2.6831
1.5	2.0122	4.0	3.0233
2.0	2.1085	4.5	3.4825
2.5	2.2444	5.0	4.1044
3.0	2.4308	5.5	4.9515

To illustrate the use of (4.14) we give the following examples.

I. Find the maximum error committed by using only 18 terms of the Rademacher series for $p(599)$. Here we have $\mu_{599}=62.777$ and $r=\mu/18=3.4876$. Hence $F(r)=2.720$ so that $|R_{18}(599)|<.396$. Actually† $R=.00027$.

II. Find $R_{21}(721)$. Here $\mu=68.8746$, $\mu/21=r=3.2797$. $F(r)=2.596$, and $|R_{21}(721)|<.341$. Actually† $R=.00041$.

We now consider the difference $d_N(n)$ between the sums of the N first terms of Rademacher series and that of Hardy and Ramanujan. That is, in view of (4.10),

$$d_N(n)=\frac{12^{1/2}}{24n-1}\sum_{k=1}^{N}A_k^*(n)\left(1+\frac{k}{\mu}\right)e^{-\mu/k}.$$

Using (4.12) we find

$$|d_N(n)| < \frac{48^{1/2}}{24n-1}\left\{\int_0^N x^{1/3}\left(1+\frac{x}{\mu}\right)e^{-\mu/x}dx + N^{1/3}\left(1+\frac{N}{\mu}\right)e^{-\mu/N}\right\}.$$

Since $e^{-\mu/x}<x/\mu$ we find, on writing $r=\lambda/N$ and $24n-1=(6\mu/\pi)^2$,

† Journal of the London Mathematical Society, vol. 11 (1936), pp. 115–116.

$$(4.15) \qquad |d_N(n)| < \frac{\pi^2}{3^{1/2}\mu}\frac{r+1}{r^2}\left\{\frac{\mu^{1/3}}{7r^{4/3}} + \frac{r^{1/3}}{3\mu}\right\}.$$

This estimate, crude though it is, shows that, for typical calculations of $p(n)$, $d_N(n)$ is sensibly zero.

We are now in a position to answer the question: When is the Hardy-Ramanujan series applicable? This may be answered in a number of ways of which the following is an example.

THEOREM 13. *If only* $2n^{1/2}/3$ *terms of the Hardy-Ramanujan series* (1.1) *be taken, the resulting sum will differ from* $p(n)$ *by less than* 1/2, *provided* $n>600$.†

Proof. If $N=2n^{1/2}/3$, then

$$(4.16) \qquad r = \frac{\mu}{N} = \frac{2\mu}{3n^{1/2}} = \frac{\pi}{9}\frac{(24n-1)^{1/2}}{n^{1/2}} < 3.847.$$

Since $n>600$, $N=\frac{2}{3}n^{1/2}>16$, and $\mu>\pi/6(24n)^{1/2}>62.832$.

$$|R_N(n)| < F(3.847)16^{-2/3} < .46.$$

Now since the right member of (4.15) is a decreasing function of μ, we obtain

$$|d_N(n)| < .0031.$$

Hence the sum of the first N terms of the Hardy-Ramanujan series differs from $p(n)$ by an amount which is less than

$$|R_N(n)| + |d_N(n)| < .46 + .0031 < 1/2$$

in absolute value.

The factor 2/3 of Theorem 13 may be made smaller by allowing the lower limit of n to increase. For example if we wish to take only $n^{1/2}/2$ terms of the series we may do so provided $n>3600$. By making a general argument we may easily prove the following:

THEOREM 14. *Let* $\delta>1$ *and let* $c=\pi(2/3)^{1/2}=2.565\cdots$. *Then* $p(n)$ *is the nearest integer to the sum of the first* $n^{1/2}/\delta$ *terms of the Hardy-Ramanujan series provided*

$$n > \frac{27^{1/2}c^6}{\delta^2}\left\{\frac{\sinh(c\delta)}{c^3\delta^3} + \frac{1}{6}\right\}^3 = O(e^{3c\delta}\delta^{-11}).$$

† The tables of $p(n)$ for $n\leqq 600$ have been published by Gupta, Proceedings of the London Mathematical Society, (2), vol. 39 (1935), pp. 148–149; vol. 42, pp. 546–549.

LEHIGH UNIVERSITY,
BETHLEHEM, PA.

A PROBLEM IN ADDITIVE NUMBER THEORY*

BY

R. D. JAMES

1. **Introduction.** Some time ago the author was asked by Professor D. N. Lehmer if there was anything known about the representation of an integer h in the form

$$h = \sum_{i=1}^{s} h_i,$$

where all the prime factors of each h_i are of a given form. A search of the literature seemed to indicate that various theorems had been conjectured but none actually proved.† For example, L. Euler stated without proof that every integer of the form $4j+2$ is a sum of two primes each of the form $4j+1$. Even the weaker statement that every integer of the form $4j+2$ is a sum of two integers which have all their prime factors of the form $4j+1$ has not yet been proved.

In view of the absence of any definite results in the literature it seems worthwhile to point out that some very interesting theorems can be obtained in an elementary way. This is done in Part I of this paper and the results are summarized in Theorems 1, 2, and 3 below. In Part II we use the method of Viggo Brun‡ to prove a general theorem and from this we deduce Theorems 4 and 5 below.

THEOREM 1. *Consider the set of all integers n_i with the property that $n_0=1$ and that every prime factor of each n_i, $i \geqq 1$ is of the form $4j+1$. Let $r=3, 4, 5$, or 6. Then every integer $N \equiv r \pmod 4$, $N \geqq r$ is a sum of exactly r integers n_i all but three of which may be taken equal to 1. Except for $r=6$ this result is the best possible in the sense that there is an infinite number of integers $N \equiv r \pmod 4$ which are not the sum of fewer than r integers n_i.*

THEOREM 2. *Let N be any integer of the form $4j+2$. If the integer $8j+2$ is of the form $2Kp_1^{2a_1} \cdots p_t^{2a_t}$, where $p_\nu \equiv 3 \pmod 4$, $\nu = 1, 2, \cdots, t$, and every prime factor of K is of the form $4j+1$, then N is a sum of exactly two integers n_i.*

* Presented to the Society, April 3, 1937; received by the editors March 23, 1937.

† L. E. Dickson, *History of the Theory of Numbers*, vol. I, Chap. XVIII, and vol. II, Chap. VIII.

‡ See the paper by H. Rademacher, Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität, vol. 3 (1924), pp. 12–30.

THEOREM 3. *Consider the set of all integers m_i with the property that $m_0=1$ and that every prime factor of m_i, $i\geqq 1$ is of the form $8j+1$ or $8j+3$. Then every odd integer $M\geqq 3$ is a sum of exactly three integers m_i and every even integer $M\geqq 4$ is a sum of exactly four integers m_i. The results for M odd and $M\equiv 0 \pmod 8$ are the best possible.*

THEOREM 4. *Every sufficiently large integer $N\equiv 2 \pmod 4$ is a sum of two integers which have all except possibly two of their prime factors of the form $4j+1$.*

THEOREM 5. *Every sufficiently large integer $M\equiv 2, 4, 6 \pmod 8$ is a sum of two integers which have all except possibly two of their prime factors of the form $8j+1$ or $8j+3$.*

PART I

2. **Preliminary lemmas.** The lemmas which follow are well known and we shall state them without proof.

LEMMA 1.* *If k is any integer such that*

$$k \not\equiv 0,\ 7,\ 12 \text{ or } 15 \pmod{16}$$

then there exist integers x_1, x_2, and x_3 such that

$$k = \sum_{\nu=1}^{3} x_\nu^2 .$$

LEMMA 2.† *If x and y have no common factor and are not both odd, every prime factor of x^2+y^2 is of the form $4j+1$.*

LEMMA 3. *If x and y have no common factor and x is odd, every prime factor of x^2+2y^2 is of the form $8j+1$ or $8j+3$.*

3. **The proof of Theorem 1.** We suppose first that $r=3$ so that N is of the form $4j+3$. We have

$$8j + 3 \not\equiv 0,\ 7,\ 12, \text{ or } 15 \pmod{16},$$

so that by Lemma 1

$$8j + 3 = \sum_{\nu=1}^{3} x_\nu^2 . \tag{1}$$

Since x_ν^2 is of the form $4j$ or $4j+1$ according as x_ν is even or odd, it follows that each x_ν in (1) must be odd. Let $x_\nu=2s_\nu+1$. Then (1) becomes

* E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, vol. I, pp. 550–555.

† Lemmas 2 and 3 follow from the fact that -1 is a quadratic residue of an odd prime p if and only if p is of the form $4j+1$, and that -2 is a quadratic residue of an odd prime p if and only if p is of the form $8j+1$ or $8j+3$.

$$8j + 3 = \sum_{\nu=1}^{3} (2s_\nu + 1)^2,$$

$$N = 4j + 3 = \sum_{\nu=1}^{3} \{s_\nu^2 + (s_\nu + 1)^2\}.$$

Obviously s_ν and $s_\nu+1$ have no common factor and are not both odd. Hence by Lemma 2 every prime factor of the integer $S_\nu^2+(S_\nu+1)^2$ is of the form $4j+1$. This proves the first part of Theorem 1 when $r=3$.

Now let $r=4$, 5, or 6. Then $N-r+3\equiv 3 \pmod 4$ and thus $N-r+3$ is a sum of exactly three integers n_i. It follows that N itself is a sum of exactly r integers n_i, all but three of which are equal to 1.

To prove the last statement of the theorem when $r=3$ or 4 we observe that since we have each $n_i\equiv 1 \pmod 4$ the congruence

$$N \equiv \sum_{j=1}^{s} n_{i_j} \pmod 4 \tag{2}$$

has no solution when $s<r$. Therefore the equation

$$N = \sum_{j=1}^{s} n_{i_j} \tag{3}$$

certainly has no solution when $s<r$.

When $r=5$ we consider the set of all integers $N=p_1^{2a_1}\cdots p_t^{2a_t}$, where every p_ν is of the form $4j+3$. It is evident that $N\equiv 1 \pmod 4$. For these integers the congruence (2) has no solution when $1<s<5$ and hence the equation (3) has no solution when $s<5$.

4. **The proof of Theorem** 2. Since the integer $8j+2$ is of the form $2Kp_1^{2a_1}\cdots p_t^{2a_t}$, where $p_\nu\equiv 3 \pmod 4$ and every prime factor of K is of the form $4j+1$, there exist integers u and v such that*

$$8j + 2 = u^2 + v^2. \tag{4}$$

By the argument used in the proof of Theorem 1, both u and v must be odd. Let $u=2y+1$, $v=2z+1$. Then (4) becomes

$$8j + 2 = 2\{y^2 + (y + 1)^2 + z^2 + (z + 1)^2\} - 2,$$

$$N = 4j + 2 = y^2 + (y + 1)^2 + z^2 + (z + 1)^2.$$

Every prime factor of $y^2+(y+1)^2$ and $z^2+(z+1)^2$ is of the form $4j+1$ and this completes the proof.

* E. Landau, loc. cit., pp. 549–550.

5. **The proof of Theorem** 3. We suppose first that $M=2k+3$. If

$$k\not\equiv 0,\ 7,\ 12, \text{ or } 15 \pmod{16}$$

we have

$$k = \sum_{i=1}^{3} x_i^2,$$

$$2k+3 = \sum_{i=1}^{3} (2x_i^2+1).$$

By Lemma 3 every prime factor of $2x_i^2+1$ is of the form $8j+1$ or $8j+3$.

If $k\equiv 0$ or 12 (mod 16) then $2k-21\equiv 3$ or 11 (mod 16). Then*

$$2k-21 = \sum_{i=1}^{3} x_i^2,$$

(5)

$$2k+3 = \sum_{i=1}^{3} (x_i^2+8).$$

In (5) every x_i is odd and the result follows from Lemma 3.

If $k\equiv 7$ or 15 (mod 16) then $2k-3\equiv 11$ (mod 16) and we have

$$2k-3 = \sum_{i=1}^{3} x_i^2,$$

$$2k+3 = \sum_{i=1}^{3} (x_i^2+2).$$

Again x_i is odd and the theorem follows as before.

The rest of the theorem is a consequence of the first part since $M-1$ is odd if M is even. The results can be shown to be the best possible by using congruential conditions similar to those used in the proof of Theorem 1.

Part II

6. **The Viggo Brun method.** In this part we use the results of the paper by Rademacher to which reference was made above. This will be cited as R.† Let $p_1, p_2, \cdots,$ be any infinite set of primes which are all distinct. Let $a_1, a_2, \cdots, b_1, b_2, \cdots,$ be any integers such that $a_i\neq b_i$. For $(\Delta, D)=1$ let

$$P(\Delta, D, x; a_1, b_1, p_1; \cdots; a_r, b_r, p_r) = P(D, x; p_1, \cdots, p_r)$$

* The case $k=0$, $M=3$ is not included here but obviously $M=3=1+1+1$.

† T. Estermann, Journal für die Reine und Angewandte Mathematik, vol. 168 (1932), pp. 106–116, has improved Rademacher's results. For the problem which we are considering, however, Estermann's method does not yield anything more.

denote the number of integers z which satisfy the conditions

$$0 < z \leqq x,\ z \equiv \Delta \pmod{D}, \qquad (z - a_i)(z - b_i) \not\equiv 0 \pmod{p_i}, \qquad (i = 1, 2, \cdots, r). \tag{6}$$

Then by R, (8) we have

$$P(D, x; p_1, p_2, \cdots, p_r) > \frac{E}{D} x - R,$$

where

$$E = 1 - 2\sum_{\alpha \leqq r} \frac{1}{p_\alpha} + 4 \sum_{\substack{\alpha \leqq r \\ \beta < \alpha}} \sum_{\beta \leqq r_1} \frac{1}{p_\alpha p_\beta} - \cdots - 2^{2n+1} \sum_{\substack{\alpha \leqq r \\ \mu < \cdots < \beta < \alpha}} \cdots \sum_{\mu < r_n} \frac{1}{p_\alpha p_\beta \cdots p_\mu},$$

$$R = (2r + 1)(2r_1 + 1)^2 \cdots (2r_n + 1)^2, \qquad r > r_1 > \cdots > r_n \geqq 1.$$

We now assume that the primes $p_1, \cdots, p_r$ are the first r primes in order of any infinite set of primes which have the property that

$$\sum_{3 \leqq p \leqq w}' \frac{1}{p} = \frac{1}{\alpha} \log \log w + c_1(\alpha) + o(1). \tag{7}$$

Here $\sum'$ or $\prod'$ denotes the sum or product over all primes of the set which are $\leqq w$. From (7) and a general theorem on infinite series* it follows that

$$\prod_{3 \leqq p \leqq w}' \left(1 - \frac{2}{p}\right) = \frac{c_2(\alpha)}{(\log w)^{2/\alpha}} + o\left(\frac{1}{(\log w)^{2/\alpha}}\right). \tag{8}$$

If $\alpha = 1$, this reduces to the case treated by Rademacher.

Now let h and h_0 be any two numbers such that

$$1 < h < h_0^\alpha, \qquad 0 < 2 \log h_0 < 1.$$

Then from (7) and (8) it follows that there is a number w_0 such that for all $w \geqq w_0$ we have

$$0 < \sum_{w < p \leqq w^h}' \frac{1}{p} < \log h_0,$$

$$\prod_{w < p \leqq w^h}' \left(1 - \frac{2}{p}\right) > \frac{1}{h_0^2} \cdot$$

These are precisely the equations (15a) which are used in R. All the results obtained there go over to the case which we are considering. Thus from R, (18) and (26) we obtain

* K. Knopp, *Theorie und Anwendung der Unendlichen Reihen*, page 218.

$$E > \prod_{\nu=1}^{r}\left(1-\frac{2}{p_\nu}\right)\left\{E_1 - h_0^2\,\Phi_2 - \frac{2h_0^4\log^6 h_0\,e^2(e^2-5)}{1-e^2h_0^2\log^2 h_0}\right\},$$

$$R < c_3 p_r^{(h+1)/(h-1)},$$

where c_3 depends only on α, h, h_0, and $E_1 > 1-2\log h_0$, $\Phi_2 < (10\log^4 h_0)/3$. If we take $h_0 = 1.3$ we find that

$$P(D, x; p_1, \cdots, p_r) > \frac{C}{D}\,\frac{x}{(\log p_r)^{2/\alpha}} - C' p_r^{(h+1)/(h-1)}, \tag{9}$$

where C and C' depend only on α, h. It is this inequality which we use to prove Theorems 4 and 5.

7. **The proof of Theorem** 4. Let x in (6) be of the form $4j+2$. Consider the infinite set of primes p which are of the form $4j+3$. In this case we have*

$$\sum_{\substack{3\leqq p\leqq w\\ p\,\equiv\,3\ (\mathrm{mod}\ 4)}} \frac{1}{p} = \frac{1}{\phi(4)}\log\log w + c_1 + o(1).$$

Then $\alpha = 2$ and we may take $h = 1.68 < (1.3)^2$. From (9) we have

$$P(D, x; p_1, \cdots, p_r) > \frac{C}{D}\,\frac{x}{\log p_r} - C' p_r^{268/68}. \tag{10}$$

Let $p_1, p_2, \cdots, p_r$ be the primes $7, 11, \cdots$ up to the largest prime of the form $4j+3$ which does not exceed $x^{1/4}$. We choose a_i and b_i in the following manner.

$$a_i = 0, \qquad b_i = x, \qquad \text{if} \quad p_i \nmid x;$$
$$a_i = 0, \qquad b_i = 1, \qquad \text{if} \quad p_i \mid x.$$

We choose Δ so that $z \equiv 1 \pmod 4$ and so that neither Δ nor $x-\Delta$ is divisible by 3. Then Δ is determined (mod 12). Using the fact that $p_r \leqq x^{1/4}$ the inequality (10) becomes

$$P(12, x; p_1, \cdots, p_r) > \frac{2C}{3}\,\frac{x}{\log x} - C' x^{268/272}.$$

Hence for x sufficiently large we have $P(12, x; p_1, \cdots, p_r) \geqq 1$. Going back to the definition of $P(12, x; p_1, \cdots, p_r)$ we see that this means that there is at least one integer z such that

$$0 < z \leqq x, \qquad z \equiv \Delta \pmod{12}, \qquad \begin{array}{ll} z(z-x) \not\equiv 0 \pmod{p_i}, & p_i \nmid x; \\ z(z-1-x) \not\equiv 0 \pmod{p_i}, & p_i \mid x. \end{array}$$

* E. Landau, loc. cit., pp. 449–450.

This shows that there is at least one integer z for which

$$x = z + (x - z),$$

where neither z nor $x-z$ is divisible by 2 or by any prime of the form $4j+3$ which does not exceed $x^{1/4}$. If a prime of the form $4j+3$ does divide z or $x-z$, then it must be greater than $x^{1/4}$. This proves that not more than three primes of the form $4j+3$ can divide z or $x-z$. The number three can be reduced to two by the following argument. Both z and $x-z$ are of the form $4j+1$, but a product of three primes of the form $4j+3$ is again of the form $4j+3$. Therefore not more than two primes of the form $4j+3$ can divide z or $x-z$. This proves Theorem 4.

8. **The proof of Theorem** 5. The proof of this theorem is only slightly different from the proof of Theorem 4. We have*

$$\sum_{\substack{3 \leq p \leq w \\ p \equiv 5 \text{ or } 7 \pmod 8}} \frac{1}{p} = \frac{2}{\phi(8)} \log \log w + 2c_1 + o(1),$$

and again $\alpha = 2$, $h = 1.68$. This time we choose Δ so that z has the following values (mod 8) and so that neither Δ nor $x-\Delta$ is divisible by 3.

$$\begin{array}{llll} z \equiv 1, & x - z \equiv 1 & \text{if} & x \equiv 2 \pmod 8, \\ z \equiv 1, & x - z \equiv 3 & \text{if} & x \equiv 4 \pmod 8, \\ z \equiv 3, & x - z \equiv 3 & \text{if} & x \equiv 6 \pmod 8. \end{array}$$

The inequality (10) then shows that not more than three primes of the form $8j+5$ or $8j+7$ can divide z or $x-z$. An argument similar to that used in the proof of Theorem 4 shows finally that not more than two primes of the form $8j+5$ or $8j+7$ can divide z or $x-z$. This completes the proof of Theorem 5.

* E. Landau, loc. cit.

The University of California,
Berkeley, Calif.

TRANSFORMATIONS OF A SURFACE BEARING A FAMILY OF ASYMPTOTIC CURVES*

BY

G. D. GORE

Introduction. It is the purpose of this paper to establish certain transformations for any non-developable surface that bears a family of asymptotic curves and is immersed in a space of n dimensions S_n ($n>3$). All surfaces mentioned hereafter will belong to S_n.

The ambient of the osculating planes at a point of a surface to all of the curves on the surface that go through the point is a space of not more than five dimensions.† The class of all surfaces in S_n for which the ambient at all points is a space of four dimensions is divided into two subclasses. One of these subclasses is composed of all surfaces in S_n that bear each a conjugate net of curves, while the other is composed of all surfaces in S_n that sustain each a family of asymptotic curves.

In the classical transformations for a surface bearing a conjugate net, the two congruences of lines tangent to the curves of the net have played basic rôles. We shall assign a similar rôle to the ∞^2 lines tangent to the asymptotic curves of a family. Although the congruence of these lines contains only a one parameter family of developable surfaces, it will be defined as a *parabolic congruence.*

To facilitate discussion, a terminology for certain geometric relations is introduced.

DEFINITION 1. *The asymptotic curves of a given family are said to be autoconjugate to the lines of a parabolic congruence if the curves of the family lie on the developable surfaces of the congruence, provided that the surface which sustains the given family is not the focal surface of the congruence.*

DEFINITION 2. *A family of asymptotic curves on a surface and a parabolic congruence, such that there is just one line of the congruence lying in each tangent plane of the surface and not passing through the point of contact, are harmonic to each other in case the developable surfaces of the parabolic congruence correspond to the curves of the family.*

* Presented to the Society, November 27, 1936; received by the editors March 16, 1937.

† Lane, *Projective Differential Geometry of Curves and Surfaces*, University of Chicago Press, 1932, p. 124.

DEFINITION 3. *If two families of asymptotic curves are autoconjugate to the same parabolic congruence, they are said to be in relation F'.*

The transformation F' of families of asymptotic lines is an analogue of the well known transformation F of conjugate systems.*

DEFINITION 4. *If the points of a surface are in a one-to-one correspondence with ∞^2 straight lines, and if corresponding points and lines are in united position, the surface is said to be transversal to the lines.*

In §1 is developed a transformation of a family of asymptotic curves which is analogous to the transformation of Levy for conjugate nets,† while in §2 is exhibited a method for determining all of the parabolic congruences autoconjugate to a given family of asymptotic curves. There is developed in §3 a method for determining a family of asymptotic curves in relation F' with a given family of curves of the same kind. A study is made in §4 of the relation of several F' transforms of a given family of asymptotic curves by means of the same parabolic congruence. The relation of two F' transforms of a given surface by means of two different parabolic congruences is considered in §5, and a *theorem of permutability* for the transformation F' is established in §6. General transversal surfaces of a parabolic congruence are examined in §7, and it is proved that these surfaces too are transformable by some of the methods which we have applied to surfaces bearing families of asymptotic curves.

1. **A transformation of a family of asymptotic curves.** Consider a non-developable surface S which sustains a family of asymptotic curves and is immersed in a projective space of n dimensions ($n>3$). A parametric vector equation of the surface may be written in terms of homogeneous coordinates as $y=y(u, v)$. We adopt the asymptotic curves as the u-curves, and any other family of curves on the surface as v-curves. The resulting coordinates y of the generating point of the surface are known to satisfy a differential equation of the form

$$y_{uu} = ay_u + by_v + cy \qquad (b \neq 0), \tag{1.1}$$

called the *point differential equation* of the surface. The surface S is any integral surface of equation (1.1).

To obtain a transformation of S, let R be a solution of (1.1). The point determined by the coordinates

$$x = Ry_u - R_u y \tag{1.2}$$

is on the line which is tangent at the point y to the u-curve of S. We shall

* See Eisenhart, *Transformations of Surfaces*, Princeton University Press, 1923, p. 34.

† See Eisenhart, loc. cit., p. 19.

show that the point x generates a surface having a family of asymptotic curves as u-curves.

By computing derivatives of (1.2) and reducing them by means of (1.1), we obtain the relations

$$
\begin{aligned}
x &= Ry_u - R_u y, \\
x_u - ax &= b(Ry_v - R_v y), \qquad (1.3)\\
x_{uu} - ax_u - bx_v - a_u x &= -2bR_v y_u + (2bR_u + b_u R)y_v - b_u R_v y.
\end{aligned}
$$

The determinant of the coefficients of y, y_u, and y_v in the right members of (1.3) is equal to zero. Hence the left members satisfy a linear relation. This relation is equivalent to the differential equation

$$
x_{uu} = \left(a + \frac{b_u}{b} + \frac{2R_u}{R}\right)x_u + bx_v + \left(a_u - \frac{2aR_u}{R} - \frac{ab_u}{b} - \frac{2bR_v}{R}\right)x, \qquad (1.4)
$$

which indicates that the u-curves of the surface $S(x)$ belong to a family of asymptotic curves.

The first of equations (1.3) indicates that the surface $S(x)$ is transversal to the parabolic congruence of lines which are tangent to the u-curves of $S(y)$. Moreover, the family of asymptotic u-curves on $S(x)$ is autoconjugate to the parabolic congruence.

From the first and second equations of (1.3), we observe that the line which is tangent to the u-curve of $S(x)$ at the point x, lies in the plane which is tangent to $S(y)$ at the point y. Hence the parabolic congruence of lines tangent to the u-curves of $S(x)$ is harmonic to the family of asymptotic u-curves on $S(y)$.

The transformation (1.2) is an analogue of the transformation of Levy for conjugate nets. Repeated application of this transformation produces a sequence of surfaces which is a close analogue of a Levy sequence of conjugate nets.

We shall now prove that all families of asymptotic curves that are autoconjugate to the parabolic congruence of lines tangent to the u-curves of $S(y)$ are obtained by transformations of the same form as (1.2).

Let the generating point of a surface transversal to the tangent lines of the u-curves of $S(y)$ have coordinates ξ. By means of a transformation $y=\theta\eta$, let new coordinates η for the generating point of $S(y)$ be chosen so that

$$
\xi = \eta_u. \qquad (1.5)
$$

This change of coordinates transforms the differential equation (1.1) into an equation of the form

$$\eta_{uu} = \alpha\eta_u + \beta\eta_v + \gamma\eta \qquad (\beta \neq 0), \tag{1.6}$$

in which certain of the coefficients are specialized by the particular choice of θ.

In order that the surface $S(\xi)$ have a family of asymptotic u-curves, it is necessary and sufficient that ξ satisfy an equation of the same form as (1.1). If, by means of equations (1.5) and (1.6) we express ξ, ξ_u, ξ_v, and ξ_{uu} in terms of η, η_u, η_v, and η_{uv}, and set the determinant of the coefficients of the latter functions equal to zero, we find that

$$\beta\gamma_u - \gamma\beta_u = 0 \tag{1.7}$$

is the only restriction on the coefficients of (1.6) in order that $S(\xi)$ have the required family of asymptotic curves.

The general solution of (1.7) is $\gamma = \beta f(v)$, a special case of which is $\gamma = f(v) = 0$. But in order for γ to be zero, the above function θ must be a solution of (1.1). Then the coordinates ξ can be written as

$$\xi = (1/\theta^2)(\theta y_u - \theta_u y).$$

If $f(v) \neq 0$, we introduce the value of γ, and verify that

$$\xi_u - \alpha\xi = \beta(\eta_v + f\eta).$$

A second transformation $\eta = \mu(v)\zeta$ is introduced, where μ satisfies the condition $d\mu/dv = -\mu f$. As a result of this transformation, the above equation and (1.5) become

$$\begin{aligned} \xi_u - \alpha\xi &= \beta\mu\zeta_v, \\ \xi &= \mu\zeta_u. \end{aligned} \tag{1.8}$$

Integrability conditions on the left members of (1.8) show that ζ satisfies a differential equation of the same form as (1.6), but with $\gamma = 0$.

The above transformations $y = \theta\eta$ and $\eta = \mu\zeta$ are equivalent to the single transformation $y = R\zeta$. Since this transformation changes (1.1) into a new equation in ζ, of the same form as (1.6), but with $\gamma = 0$, the function R is a solution of (1.1).

The second of equations (1.8) can be written in the form $\xi = (\mu/R^2)(Ry_u - R_u y)$. This establishes

THEOREM 1. *Let $S(y)$ be a surface bearing a family of asymptotic u-curves, and for which the point differential equation is* (1.1). *Let $S(x)$ be any surface which sustains a family of asymptotic curves autoconjugate to the parabolic congruence of lines tangent to the u-curves of $S(y)$. Then the transformation which sends $S(y)$ into $S(x)$ may be represented by a relation of the form $x = Ry_u - R_u y$, in which R is a solution of* (1.1).

We make use of a second solution R' of equation (1.1) to construct the transformation of $S(x)$ which is represented by the equation

$$x' = R'y_u - R'_u y. \tag{1.9}$$

From equations (1.1) and (1.9) is derived the relation

$$x'_u - ax' = b(R'y_v - R'_v y), \tag{1.10}$$

which is similar to the second of (1.3). Equations (1.9) and (1.10) together with (1.3) indicate that the line tangent at the point x' to the u-curve of $S(x')$, and the corresponding line tangent to the u-curve of $S(x)$, lie in the plane which is tangent at the point y to $S(y)$. Since these lines lie in a plane, they intersect in a point having coordinates x''. We shall prove that the point x'' generates a surface which has a family of asymptotic lines as parametric u-curves, and that this family of curves is autoconjugate to each of the parabolic congruences formed by the lines tangent to the u-curves of $S(x)$ and $S(x')$ respectively.

Since R' is a solution of (1.1), the function

$$\theta = RR'_u - R'R_u \tag{1.11}$$

is a solution of the point differential equation of $S(x)$. This fact follows from replacing x by R' in (1.2). For a similar reason, the function

$$\theta' = R'R_u - RR'_u \tag{1.12}$$

is a solution of the point differential equation of $S(x')$. We define $\psi = RR'_v - R_vR'$. Then by (1.11) and (1.1),

$$\begin{aligned} \theta_u &= a\theta + b\psi, \\ \theta_u &= a\theta' - b\psi. \end{aligned} \tag{1.13}$$

To determine the coordinates of the point of intersection of the two corresponding tangent lines to the u-curves of $S(x)$ and $S(x')$, we list the following relations

$$\begin{aligned} x &= Ry_u - R_u y, \\ x_u &= R(ay_u + by_v) - (aR_u + bR_v)y, \\ x' &= R'y_u - R'_u y, \\ x'_u &= R'(ay_u + by_v) - a(R'_u + bR'_v)y. \end{aligned} \tag{1.14}$$

By eliminating y, y_u, and y_v from (1.14), we obtain the relation

$$R'[\theta x_u - (a\theta + b\psi)x] = -R[\theta' x'_u - (a\theta' - b\psi)x'].$$

After using equations (1.13) to reduce the coefficients of x and x' in the above equation, we have the coordinates

$$x'' = R'(\theta x_u - \theta_u x) = - R(\theta' x'_u - \theta'_u x').$$

Since θ and θ' are solutions of the point differential equations of $S(x)$ and $S(x')$ respectively, the surface $S(x'')$ is by (1.2) a transform of $S(x)$ and $S(x')$ alike. By virtue of these transformations, $S(x'')$ has a family of asymptotic curves as u-curves, and this family is autoconjugate to each of the two parabolic congruences formed by lines tangent to the u-curves of the surfaces $S(x)$ and $S(x')$ respectively.

Since, by the remark following equation (1.10), the line tangent at the point x'' to the u-curve of $S(x'')$ lies in the plane tangent at the point x to $S(x)$, and lies also in the plane tangent at the point x' to $S(x')$, it follows that the families of asymptotic curves of both $S(x)$ and $S(x')$ are harmonic to the parabolic congruence formed by the totality of these tangent lines. From these results, we state

THEOREM 2. *If two families of asymptotic curves are autoconjugate to the same parabolic congruence, they are both harmonic to a second parabolic congruence.*

2. **Parabolic congruences autoconjugate to a family of asymptotic curves.** Let $S(x)$ be a surface having a family of asymptotic u-curves. Denote by $S(y)$ the focal surface of a parabolic congruence that is autoconjugate to the family of asymptotic curves on $S(x)$.

The coordinates x satisfy a differential equation

$$x_{uu} = a x_u + b x_v + c x, \tag{2.1}$$

and the coordinates y satisfy an equation

$$y_{uu} = A y_u + B y_v + C y. \tag{2.2}$$

Since the point x is on the line tangent at point y to the u-curve of $S(y)$, the coordinates x and y satisfy a relation

$$y_u + \lambda y = \mu x. \tag{2.3}$$

To obtain the consequences of equations (2.1), (2.2), and (2.3), we take the derivative with respect to u of (2.3), and reduce the result by (2.2) and (2.3). The relation reduces to the equation

$$B y_v + (C + \lambda_u - \lambda^2 - A\lambda) y = \mu x_u + (\mu_u - \mu A - \mu\lambda) x,$$

which can be written in an abbreviated form with (2.3) to give the system

$$
(2.4)\qquad
\begin{aligned}
y_u + \alpha y &= \beta x,\\
y_v + \rho y &= \sigma x_u + \tau x.
\end{aligned}
$$

Equations (2.4) imply relations of the form (2.2) and (2.3). To demonstrate this fact, we compute the derivative with respect to u of the first equation of (2.4) and eliminate x.

Henceforth we shall investigate the consequences of equations (2.4) in view of (2.1). On differentiating (2.4) we obtain the system

$$
(2.5)\qquad
\begin{aligned}
y_u + \alpha y &= \beta x,\\
y_{uv} + \alpha y_v + \alpha_v y &= \beta x_v + \beta_v x,\\
y_v + \rho y &= \sigma x_u + \tau x,\\
y_{uv} + \rho y_u + \rho_u y &= (\sigma_u + a\sigma + \tau)x_u + \sigma b x_v + (\tau_u + \sigma c)x.
\end{aligned}
$$

The right members of these equations are linearly dependent. Unless the determinant of the coefficients of the left members vanishes there exists a linear relation in the functions y_{uv}, y_u, y_v, and y. Such a relation together with (2.2) would restrict the surface $S(y)$ to be developable. We consider the case for which $S(y)$ is not developable, and for which the determinant vanishes. The expanded form of the determinant set equal to zero is

$$
(2.6)\qquad \rho_u - \alpha_v = 0.
$$

This equation makes it possible, by means of a transformation $y=\phi\bar{y}$, to reduce equations (2.4) to the form

$$
(2.7)\qquad
\begin{aligned}
y_u &= gx,\\
y_v &= Rx_u + Sx.
\end{aligned}
$$

On applying integrability conditions to the left members of (2.7), we obtain the relation

$$
x_u(Ra + R_u + S) + x_v(Rb - g) + x(S_u + Rc - g_v) = 0.
$$

The left member of this equation must vanish identically in x and its derivatives. The coefficients set equal to zero give

$$
(2.8)\qquad
\begin{aligned}
S &= -(R_u + Ra),\\
g &= Rb,\\
S_u &= g_v - Rc.
\end{aligned}
$$

By using the first and second equations of (2.8) to eliminate g and S from the third, we show that R is a solution of the equation

$$
(2.9)\qquad R_{uu} = -aR_u - bR_v + (c - a_u - b_v)R.
$$

These results are summarized in the following theorem.

THEOREM 3. *Let there be given an integral surface $S(x)$ of the point differential equation* (2.1), *and let R be a solution of* (2.9), *with g and S determined by the first two of equations* (2.8). *Then the coordinates y obtained by quadratures from* (2.7) *determine the generating point of the focal surface of a parabolic congruence that is autoconjugate to the family of asymptotic u-curves on $S(x)$.*

3. **Families of asymptotic curves in relation F'.** Two families of asymptotic curves which are autoconjugate to the same parabolic congruence are, by Definition 3, in relation F'.

Let $S(x)$ and $S(x')$ be two surfaces bearing each a family of asymptotic curves, and so related that the two families are in relation F'. Let $S(y)$ be the focal surface of the parabolic congruence to which they are autoconjugate, the point differential equation of this surface being

$$y_{uu} = \alpha y_u + \beta y_v + \gamma y. \tag{3.1}$$

By reason of Theorem 1, the coordinates x and x' can be chosen so as to satisfy the following relations

$$\begin{aligned} x &= Ry_u - R_u y, \\ x_u - \alpha x &= \beta(Ry_v - R_v y), \\ x' &= R'y_u - R_u' y, \\ x_u' - \alpha x' &= \beta(R'y_v - R_v' y), \end{aligned} \tag{3.2}$$

in which R and R' are solutions of (3.1) for which $RR_u' - R'R_u \neq 0$.

As a result of eliminating y and its derivatives from (3.2), we obtain the following relations:

$$\begin{aligned} x_u - Ax &= Mx_u' + Nx', \\ x_v - Bx &= Px_u' + Mx_v' + Qx'. \end{aligned} \tag{3.3}$$

The coordinates x satisfy an equation of the form

$$x_{uu} = ax_u + bx_v + cx, \tag{3.4}$$

and the coordinates x' satisfy a similar one. By computing the derivative with respect to v of the first of (3.3), and the derivative with respect to u of the second, and reducing the resulting system by (3.4), it can be shown that unless the determinant $A_v - B_u$ is equal to zero, the point x lies in the plane tangent to $S(x')$ at the point x'.

With the condition $A_v - B_u = 0$, equations (3.3) are transformed by the substitution $x = \theta\bar{x}$ into the form

$$(3.5)\qquad \begin{aligned} \bar{x}_u &= Mx_u' + Nx', \\ \bar{x}_v &= Px_u' + Mx_v' + Qx'. \end{aligned}$$

On taking suitable linear combinations of (3.5), we obtain the relations

$$\begin{aligned} x_u' - A'x' &= \overline{M}\bar{x}_u, \\ x_v' - B'x' &= \overline{P}\bar{x}_u + \overline{M}\bar{x}_v. \end{aligned}$$

An argument similar to the above suffices to show that these equations can be transformed, by a substitution $x' = \theta'\bar{x}'$, so that $\overline{A}' = \overline{B}' = 0$. We therefore write

$$(3.6)\qquad \begin{aligned} \bar{x}_u' &= m\bar{x}_u, \\ \bar{x}_v' &= p\bar{x}_u + m\bar{x}_v. \end{aligned}$$

Integrability conditions applied to the left members of (3.6) lead to a differential equation in $\bar{x}$ of the form

$$(3.7)\qquad \bar{x}_{uu} = \bar{a}\bar{x}_u + \bar{b}\bar{x}_v,$$

where

$$(3.8)\qquad \begin{aligned} m_u &= -p\bar{b}, \\ m_v &= p_u + \bar{a}p. \end{aligned}$$

We observe that in order to reduce (3.4) to the form (3.7) by a transformation $x = \theta\bar{x}$, it is necessary and sufficient that θ be a solution of (3.4). Coefficients of (3.7) are $\bar{a} = a - 2\theta_u/\theta$, $\bar{b} = b$.

By eliminating m from (3.8) it is found that p satisfies the equation

$$(3.9)\qquad p_{uu} + \bar{a}p_u + \bar{b}p_v + (\bar{a}_u + \bar{b}_v)p = 0.$$

We have in conclusion

THEOREM 4. *Let $S(x)$ be an integral surface of* (3.4), *and let θ, p, and m be solutions of* (3.4), (3.9), *and* (3.8) *respectively. Then the coordinates x' obtained by quadratures from the equations*

$$(3.10)\qquad \begin{aligned} x_u' &= m\left(\frac{x}{\theta}\right)_u, \\ x_v' &= p\left(\frac{x}{\theta}\right)_u + m\left(\frac{x}{\theta}\right)_v, \end{aligned}$$

determine the generating point of a surface $S(x')$ having a family of asymptotic u-curves in relation F' with the family of asymptotic u-curves on $S(x)$.

One readily verifies that the point having coordinates $\eta = x' - m(x/\theta)$ is

the generating point of the focal surface of the parabolic congruence autoconjugate to the asymptotic curves on $S(x)$ and $S(x')$.

4. **Transformations F' with the same parabolic congruence.** Let $S(x)$ and $S(x')$ be two surfaces with families of asymptotic u-curves in relation F', as represented by the equations

$$x'_u = mx_u, \qquad x'_v = px_u + mx_v. \tag{4.1}$$

The coordinates x satisfy a differential equation

$$x_{uu} = ax_u + bx_v. \tag{4.2}$$

In order to determine on a third surface a family of asymptotic curves that is autoconjugate to the parabolic congruence of lines which connect corresponding points of $S(x)$ and $S(x')$, we make the change of coordinates

$$x' = \theta' x'_1 \tag{4.3}$$

and define the coordinates of a point on the line xx' as

$$x_1 = x - \theta x'_1 .$$

In terms of the old coördinates

$$x_1 = x - \frac{\theta}{\theta'} x'. \tag{4.4}$$

The functions θ and θ' will be determined so that x_1 and x'_1 satisfy a system of equations similar to (4.1).

By differentiating (4.3) and (4.4), and applying (4.1), we establish the equations

$$\begin{aligned} x_{1u} &= \frac{1}{m}\left[x'_{1u}(\theta' - m\theta) + x'_1(\theta'_u - m\theta_u)\right], \\ x_{1v} &= \frac{1}{m}\left[x'_{1v}(\theta' - \theta)\, m - x'_{1u}\frac{p\theta'}{m} + x'_1\left(\theta'_v - m\theta_v - \frac{p}{m}\theta'_u\right)\right]. \end{aligned} \tag{4.5}$$

If these equations are to take the form of (4.1), the coefficients of x'_1 must vanish. This gives the conditions

$$\theta'_u = m\theta_u, \qquad \theta'_v = p\theta_u + m\theta_v. \tag{4.6}$$

From the integrability conditions on (4.6), θ is a solution of (4.2), and θ' is obtained by quadratures on (4.6). The function θ' is a solution of the point

differential equation of $S(x')$. It can also be verified that

$$
(4.7)\qquad
\begin{aligned}
\frac{\partial}{\partial u}\left(\frac{x_1\theta'}{\theta}\right) &= (\theta' - m\theta)\frac{\partial}{\partial u}\left(\frac{x}{\theta}\right),\\
\frac{\partial}{\partial v}\left(\frac{x_1\theta'}{\theta}\right) &= (\theta' - m\theta)\frac{\partial}{\partial v}\left(\frac{x}{\theta}\right) - p\theta\frac{\partial}{\partial u}\left(\frac{x}{\theta}\right).
\end{aligned}
$$

The above deductions justify the following theorem:

THEOREM 5. *Let there be given two surfaces $S(x)$ and $S(x')$ sustaining families of asymptotic curves in relation F', and let the coordinates x and x' be chosen so that equations* (4.1) *and* (4.2) *hold. Then in terms of θ, a solution of* (4.2), *and θ', a corresponding solution of* (4.6), *the coordinates x_1 defined by* (4.4) *determine the generating point of a surface $S(x_1)$ which has a family of asymptotic u-curves in relation F' with the corresponding families on $S(x)$ and $S(x')$.*

As an example of the above transformation, we shall establish the following theorem:

THEOREM 6. *If a surface $S(x)$ bearing a family of asymptotic curves lies on a hyperquadric, any parabolic congruence autoconjugate to the family meets the hyperquadric again in a surface bearing a family of asymptotic curves in relation F' with the first family.*

Let the equation of the hyperquadric be written as

$$
(4.8)\qquad \sum a_{ik}x^i x^k = 0.
$$

Let the chosen parabolic congruence be autoconjugate to the family of asymptotic curves on $S(x')$, where the coordinates x' of the generating point of $S(x')$ are defined by (4.1). From differentiating (4.8) and reducing the results by (4.2), we obtain the relations

$$
\sum a_{ik}(x_u^i x^k + x^i x_u^k) = \sum a_{ik}(x_v^i x^k + x^i x_v^k) = \sum a_{ik} x_u^i x_u^k = 0.
$$

By means of these relations, it can be shown that the function

$$
(4.9)\qquad \theta = \sum a_{ik}(x^i x'^k + x'^i x^k)
$$

is a solution of the point equation (4.2) of $S(x)$, and that the function

$$
(4.10)\qquad \theta' = \sum a_{ik} x'^i x'^k
$$

is a corresponding solution of (4.6).

If $S(x_1)$ is determined by the transformation (4.4), using θ and θ' from (4.9) and (4.10), it is easy to show that x_1 satisfies the equation (4.8) of the hyperquadric. The surfaces $S(x_1)$ and $S(x)$ are in relation F' by transformation (4.4).

As a second example of the transformation, we consider

THEOREM 7. *If a transversal surface of a parabolic congruence lies in a hyperplane, it has on it a family of asymptotic curves that is autoconjugate to the parabolic congruence.*

Let the given surface be transversal to the parabolic congruence of lines that join corresponding points of $S(x)$ and $S(x')$, and define the coordinates of its generating point as

$$\xi = x - \lambda x'.$$

The equation of the hyperplane may be taken as $\xi^i = 0$. If λ is determined so that $\xi^i = 0$, its value is $\lambda = x^i/x'^i$. As a consequence

$$\xi = x - \frac{x^i}{x'^i} x'. \tag{4.11}$$

The functions x^i and x'^i have the properties of θ and θ' which are required by the F' transformation (4.4). Hence $S(\xi)$ is an F' transform of each of $S(x)$ and $S(x')$.

For the F' transformation (4.4), which sends $S(x)$ into $S(x_1)$ by means of the auxiliary surface $S(x') \equiv S(x_1')$, we wish to determine an inverse. That is, we wish to determine a pair of functions θ^{-1} and $(\theta')^{-1}$ such that

$$x = x_1 - \frac{\theta^{-1}}{(\theta')^{-1}} x_1', \tag{4.12}$$

where θ^{-1} is a solution of the point differential equation of $S(x_1)$, and $(\theta')^{-1}$ is related to θ^{-1} by equations similar to (4.6) which give the relations between θ' and θ. Analogues of (4.6) are obtained by solving (4.5), in view of (4.6), for x_{1u}' and x_{1v}', then replacing x_1' by $(\theta')^{-1}$ and x_1 by θ^{-1}. The equations are

$$\begin{aligned} (\theta')_u^{-1} &= \frac{m}{\theta' - m\theta} \theta_u^{-1}, \\ (\theta')_v^{-1} &= \frac{p\theta'}{(\theta' - m\theta)^2} \theta_u^{-1} + \frac{m}{\theta' - m\theta} \theta_v^{-1}. \end{aligned} \tag{4.13}$$

Equations (4.13), as well as the point differential equations of $S(x_1)$ and $S(x_1')$ obtainable from them by integrability conditions, are satisfied by the functions

$$\theta^{-1} = -\frac{\theta}{\theta'}, \qquad (\theta')^{-1} = \frac{1}{\theta'}. \tag{4.14}$$

These functions also give a result consistent with (4.4) when they are substituted into (4.12). It can be verified that

$$\begin{aligned} \frac{\partial}{\partial u}\left(\frac{1}{\theta^{-1}}\right) &= (m\theta - \theta')\frac{\partial}{\partial u}\left(\frac{1}{\theta}\right), \\ \frac{\partial}{\partial v}\left(\frac{1}{\theta^{-1}}\right) &= (m\theta - \theta')\frac{\partial}{\partial v}\left(\frac{1}{\theta}\right) + p\theta\frac{\partial}{\partial u}\left(\frac{1}{\theta}\right). \end{aligned} \tag{4.15}$$

THEOREM 8. *If $S(x)$ is transformed into $S(x_1)$ by means of θ, θ', and the auxiliary surface $S(x')$, as expressed in* (4.4), *then $S(x_1)$ is transformed into $S(x)$ by means of the functions θ^{-1}, $(\theta')^{-1}$, and x_1', where θ^{-1}, $(\theta')^{-1}$, and x_1' are defined by* (4.14) *and* (4.3), *and the transformation is expressed by* (4.12).

If by means of a second solution θ_2 of the point equation (4.2), $S(x)$ is transformed into a second surface $S(x_2)$ by the relation

$$x_2 = x - \frac{\theta_2}{\theta_2'}x', \tag{4.16}$$

the surfaces $S(x_1)$ and $S(x_2)$ are in relation F'. It is useful to know by what analytical relation $S(x_1)$ is transformed into $S(x_2)$. By subtracting equation (4.4) from (4.16) we obtain

$$x_2 = x_1 - \frac{\theta_3 x'}{\theta_3'\theta'}, \tag{4.17}$$

where the function

$$\theta_3 = \theta_2 - \frac{\theta}{\theta'}\theta_2' \tag{4.18}$$

is by (4.4) a solution of the point equation of $S(x_1)$, and the function

$$\theta_3' = \frac{\theta_2'}{\theta'} \tag{4.19}$$

is a solution of the point equation of $S(x'/\theta')$.

5. **Transformations F' by two congruences.** Let $S(x')$ and $S(x'')$ be two F' transforms of $S(x)$, where the differential equations of the transformations are

$$\begin{aligned} x_u' &= m_1 x_u, & x_v' &= p_1 x_u + m_1 x_v; \\ x_u'' &= m_2 x_u, & x_v'' &= p_2 x_u + m_2 x_v, \end{aligned} \tag{5.1}$$

in which the coordinates x satisfy (4.2). By means of $S(x')$, $S(x'')$, and a

solution θ_1 of (4.2) we obtain two more F' transforms of $S(x)$, which are represented by the equations

$$x_{1,1} = x - \frac{\theta_1}{\theta_1'} x', \qquad x_{1,2} = x - \frac{\theta_1}{\theta_1''} x'', \tag{5.2}$$

in which θ_1' and θ_1'' are obtained from (5.1) by replacing x by θ throughout.

There is an F' transform of $S(x'')$ by means of θ_1'' and $S(x')$. The coordinates of its generating point are determined by the equation

$$x_{1,1}''' = x'' - \frac{\theta_1''}{\theta_1'} x'. \tag{5.3}$$

It is easy to show that the point $x_{1,1}'''$ is the intersection of the lines $x'x''$ and $x_{1,1}x_{1,2}$.

By differentiating (5.3) and the first of (5.2), we can establish, at the end of some labor, the following equations:

$$\begin{aligned} \frac{\partial}{\partial u}(x_{1,1}''') &= \frac{m_2\theta_1' - m_1\theta_1''}{\theta_1' - m_1\theta_1} \frac{\partial}{\partial u}(x_{1,1}), \\ \frac{\partial}{\partial v}(x_{1,1}''') &= \left[\frac{m_2\theta_1' - m_1\theta_1''}{(\theta_1' - m_1\theta_1)^2} p_1\theta_1 + \frac{p_2\theta_1' - p_1\theta_1''}{\theta_1' - m_1\theta_1}\right] \frac{\partial}{\partial u}(x_{1,1}) \\ &\quad + \frac{m_2\theta_1' - m_1\theta_1''}{\theta_1' - m_1\theta_1} \frac{\partial}{\partial v}(x_{1,1}). \end{aligned} \tag{5.4}$$

These equations show that $S(x_{1,1}''')$ and $S(x_{1,1})$ are in relation F'. It follows that $S(x_{1,1}''')$ and $S(x_{1,2})$ are in relation F'.

An inverse of the F' transformation (4.4) is given by (4.12) in view of (4.14). On adapting these equations to the first of (5.2) we determine $-\theta_1/\theta_1'$ as a solution of the point differential equation of $S(x_{1,1})$. If we set $x_{1,1} = -\theta_1/\theta_1'$ in (5.4), there is obtainable by quadrature a corresponding solution of the point equation of $S(x_{1,1}''')$. We verify that $x_{1,1}''' = -\theta_1''/\theta_1'$ is such a solution. Using these two solutions, we construct the transformation

$$x_{1,2}' = x_{1,1} - \frac{\theta_1}{\theta_1''} x_{1,1}'''. \tag{5.5}$$

It can be shown that $x_{1,2}' \equiv x_{1,2}$ by means of equations (5.2) and (5.3). These results are summarized in

THEOREM 9. *If a family of asymptotic curves on $S(x)$ is transformed into two other families of asymptotic curves on $S(x_{1,1})$ and $S(x_{1,2})$ respectively by means of the same function θ_1, the latter two families are in relation F'; moreover,*

any two of the three families $S(x)$, $S(x_{1,1})$, $S(x_{1,2})$ are transforms of the third by means of the same solution of the point differential equation of the third.

The three families form a close analogy to a triad of conjuage nets.*

6. **A theorem of permutability of transformation F'.** From (5.2) it can be seen that a solution of the point differential equation of $S(x_{1,1})$ is given by

$$\theta_{12} = \theta_2 - \frac{\theta_1}{\theta_1'}\theta_2', \tag{6.1}$$

where θ_2 is a solution of (4.2), and θ_2' is a corresponding solution of the first pair of (5.1). To get a solution of the point equation of $S(x_{1,1}''')$, let θ_2'' be a solution of the second pair of (5.1) corresponding to $x=\theta_2$; then from (5.3) we have the desired solution

$$\theta_{12}''' = \theta_2'' - \frac{\theta_1''}{\theta_1'}\theta_2'. \tag{6.2}$$

By means of the above solutions θ_{12} and θ_{12}''', we construct the following transformation of $S(x_{1,1})$:

$$x_{(12)} = x_{1,1} - \frac{\theta_2\theta_1' - \theta_1\theta_2'}{\theta_2''\theta_1' - \theta_1''\theta_2'}x_{1,1}'''. \tag{6.3}$$

Using the solution θ_2, and equations similar to (5.2) we have two more transformations of $S(x)$ as follows:

$$x_{2,1} = x - \frac{\theta_2}{\theta_2'}x', \qquad x_{2,2} = x - \frac{\theta_2}{\theta_2''}x''. \tag{6.4}$$

Corresponding points of $S(x)$, $S(x_{1,2})$, and $S(x_{2,2})$ are on a straight line. They are in relation F' by pairs. To determine the analytic relation by which $S(x_{1,2})$ is transformed into $S(x_{2,2})$ we compare equations (5.2) and (6.4) to (4.4) and (4.16), and draw conclusions corresponding to (4.17), (4.18), and (4.19). The results show that

$$\theta_2 - \frac{\theta_1}{\theta_1''}\theta_2'' \tag{6.5}$$

is the required solution of the point equation of $S(x_{1,2})$, and that θ_2''/θ_1'' is a solution of the point equation of $S(x''/\theta_1'')$. That is, $S(x_{2,2})$ is a transform of $S(x_{1,2})$ by means of the solution (6.5).

From equations (5.5), placing $x_{1,2}' = x_{1,2}$, and (6.3), it is seen that $x_{(12)}$ is a transform of $S(x_{1,2})$ by means of the function

* Eisenhart, loc. cit., p. 44.

$$(6.6) \qquad \theta_{12} - \frac{\theta_1}{\theta_1''} \theta_{12}''',$$

which is a solution of the point equation of $S(x_{1,2})$. This expression is reducible to (6.5) by (6.1) and (6.2). Hence we have shown that $S(x_{2,2})$ and $S(x_{(12)})$ are transforms of $S(x_{1,2})$ by means of the same solution of its point differential equation. It follows that $S(x_{(12)})$ and $S(x_{2,2})$ are in relation F'. We shall now investigate $S(x_{(12)})$ through its relation to $S(x_{2,2})$ and $S(x_{2,1})$.

The point where the line $x'x''$ intersects the line $x_{2,2}x_{2,1}$ has coordinates

$$(6.7) \qquad x_{2,2}'''' = x' - \frac{\theta_2'}{\theta_2''} x'',$$

which are obtained by subtracting equations (6.4). From equations (6.4) we obtain as a solution of the point differential equation of $S(x_{2,2})$,

$$\theta_{21} = \theta_1 - \frac{\theta_2}{\theta_2''} \theta_1' .$$

A corresponding solution θ_{21}'''' is obtained from (6.7) as

$$\theta_{21}'''' = \theta_1' - \frac{\theta_2'}{\theta_2''} \theta_1''.$$

From these two solutions, we construct the transformation

$$x_{(21)} = x_{2,2} - \frac{\theta_1\theta_2'' - \theta_2\theta_1''}{\theta_1'\theta_2'' - \theta_2'\theta_1''} x_{2,2}'''' .$$

Using the foregoing equations, it is easy to show that

$$x_{(12)} - x = x_{(21)} - x = \frac{(\theta_1''\theta_2 - \theta_2''\theta_1)x' + (\theta_2'\theta_1 - \theta_1'\theta_2)x''}{\theta_1'\theta_2'' - \theta_1''\theta_2'} .$$

In order to estimate the prevalence of the transformations F' that exist for given families of asymptotic curves, we count the constants in the foregoing quadratures. From the manner in which θ_1', θ_1'', θ_2', θ_2'' were obtained, each contains an arbitrary constant. If $S(x_{1,1})$ and $S(x_{2,2})$ are chosen transforms of $S(x)$, the constants in θ_1' and θ_2'' are determined by the choice. The constants in θ_2' and θ_2'' are left arbitrary. From these facts, we state for transformations F' a theorem of permutability.

THEOREM 10. *If $S(x_{1,1})$ and $S(x_{2,2})$ are two F' transforms of $S(x)$ by means of functions θ_1 and θ_2, and two distinct parabolic congruences, there exist ∞^2 surfaces $S(x_{(12)})$, each bearing a family of asymptotic curves in relation F' with $S(x_{1,1})$ and $S(x_{2,2})$.*

By employing the notation used in §§4, 5, and 6, which is similar to that used by Eisenhart,* the equations in these sections can be given metric interpretations which yield a theory of parallel transformations for families of asymptotic curves, and also radial transformations of the same.

7. **General transversal surfaces of a parabolic congruence.** Consider $S(x)$, the most general transversal surface of a parabolic congruence. Let the curves cut out on $S(x)$ by the developable surfaces of the congruence be used as parametric u-curves. It has been demonstrated by the author† that under these conditions the coordinates x satisfy an integrable system of differential equations of the form

$$(7.1)\qquad \begin{aligned} x_{uuu} &= ax_{uu} + 2bx_{uv} + cx_{vv} + dx_u + ex_v + fx, \\ x_{uuv} &= a'x_{uu} + 2b'x_{uv} + c'x_{vv} + d'x_u + e'x_v + f'x. \end{aligned}$$

These equations require the following conditions of integrability:

$$(7.2)\qquad \begin{aligned} &c = 2b - c' = 0, \\ &ac' - c_u' + e - 2b'c' = 0, \\ &2b_v + 2ab' - 2a'b + d - 2b_u' - 4b'^2 - e' = 0, \\ &a_v - a_u' - d' - 2a'b' = 0, \\ &e_v + ae' - a'e - 2b'e' - e_u' + f = 0, \\ &d_v + ad' - a'd - 2b'd' - d_u' - f' = 0, \\ &f_v + af' - a'f - 2b'f' - f_u' = 0. \end{aligned}$$

In the sense of Definition 1 the u-curves of the surface $S(x)$ are autoconjugate to the parabolic congruence to which $S(x)$ is transversal.

To determine the focal surface $S(y)$ of a parabolic congruence which is autoconjugate to the u-curves of $S(x)$, an integral surface of (7.1), consider the point having coordinates

$$(7.3)\qquad y = x_{uu} - 2b'x_u - c'x_v + (c_v' - a'c' - e')x.$$

Using equations (7.1) and (7.2) it can be verified that x and y are related by equations of the form

$$(7.4)\qquad \begin{aligned} y_u - (a - 2b')y &= Gx, \\ y_v - a'y &= hx + kx_u. \end{aligned}$$

On eliminating x from (7.4) it is found that the coordinates y satisfy an equation

* Eisenhart, loc. cit., ch. 2.

† Gore, University of Chicago Dissertation, 1932, p. 47.

$$y_{uu} = \alpha y_u + \beta y_v + \gamma y. \tag{7.5}$$

The first of equations (7.4) shows that $S(x)$ is transversal to the u-tangent lines of $S(y)$, while equation (7.5) indicates that the u-curves of $S(y)$ are a family of asymptotic curves.

Since the surface $S(y)$ can be transformed into a sequence of surfaces by transformations of the type of (1.2) and since the surface $S(x)$ is transversal to the connecting tangent lines between two consecutive surfaces, we can make use of a general inscribing theorem* which we quote:

"*Let T denote a sequence of surfaces in which the points of each surface Σ_{i+1} are joined in a one-to-one manner to the corresponding points of Σ_i by a set Ω_i of ∞^2 osculating spaces of ν dimensions belonging to the curves on the surface Σ_i. Let Σ_r' be any surface that is transversal to the set of osculants Ω_r. Then it follows that the transversal surface Σ_r' belongs to a sequence of surfaces T' which is inscribed in the given sequence T.*"

The above theorem shows that the surface $S(x)$ belongs to a sequence that is inscribed in the sequence to which $S(y)$ belongs. But since $S(y)$ can be transformed into a multiplicity of sequences, the same is true of $S(x)$. The transformations that operate to produce them are similar to (1.2).

---

* Gore, *Inscribed sequences of surfaces associated with generalized sequences of Laplace*, these Transactions, vol. 36 (1934), p. 532.

CENTRAL Y.M.C.A. COLLEGE,
CHICAGO, ILL.

ON FUNCTIONS WITH BOUNDED DERIVATIVES*

BY

OYSTEIN ORE

1. The following well known theorem is due to A. Markoff:†

Let $f_n(x)$ be a polynomial of degree n and let M_0 be the maximum of $|f_n(x)|$ in the interval (a, b). One then has for the same interval

$$(1) \qquad |f_n'(x)| \leqq \frac{2M_0 \cdot n^2}{b-a}.$$

The equality sign can only hold for the polynomials

$$(2) \qquad f_n(x) = c \cdot T_n\left(\frac{2x-a-b}{b-a}\right),$$

where $T_n(x)$ is the nth *Tschebyschef* polynomial.

We shall show that this theorem may be formulated in such a manner that it holds for arbitrary functions with a certain number of derivatives. A polynomial of degree n is characterized by the property that its $(n+1)$st derivative vanishes identically. The theorem of Markoff may be considered as a theorem on functions having a bounded $(n+1)$st derivative in a certain interval. One also obtains bounds for all derivatives from the first to the nth. Similar results may be obtained for analytic functions bounded together with some derivative in a part of the complex plane. The proofs are simple and depend upon the polynomial character of the Taylor expansion. It should be remarked that the same extension principle may be applied to several other theorems on polynomials.

2. We shall first prove:

THEOREM 1. *Let $f(x)$ be a function for which derivatives up to the $(n+1)$st exist. Let*

$$(3) \qquad |f(x)| \leqq M_0, \qquad |f^{(n+1)}(x)| \leqq M_{n+1}$$

in the interval (a, b). Then one has in the same interval

$$(4) \qquad |f'(x)| \leqq \frac{2n^2}{b-a}\cdot\left(M_0 + M_{n+1}\cdot\frac{(b-a)^{n+1}}{(n+1)!}\right).$$

* Presented to the Society, March 26, 1937; received by the editors March 20, 1937.

† A. Markoff, *Sur une question posée par Mendeleieff*, Bulletin of the Academy of Sciences of St. Petersburg, vol. 62 (1889), pp. 1–24.

To prove Theorem 1 we apply the Taylor expansion in the following form

$$f(x+h) = f(x) + h\frac{f'(x)}{1!} + \cdots + h^n\frac{f^{(n)}(x)}{n!} + R_n, \tag{5}$$

where*

$$R_n = (-1)^n\frac{1}{n!}\int_x^{x+h}(t-x)^n\cdot f^{(n+1)}(t)dt. \tag{6}$$

We now suppose x fixed in the interval (a, b) and let h vary such that $x+h$ belongs to the same interval. Hence

$$a - x \leqq h \leqq b - x. \tag{7}$$

For the remainder term (6) we then easily find

$$|R_n| \leqq M_{n+1}\frac{|h|^{n+1}}{(n+1)!} \leqq M_{n+1}\cdot\frac{(b-a)^{n+1}}{(n+1)!}. \tag{8}$$

We next consider the polynomial

$$P(h) = f(x) + h\frac{f'(x)}{1!} + \cdots + h^n\cdot\frac{f^{(n)}(x)}{n!}. \tag{9}$$

From (5) follows

$$P(h) = f(x+h) - R_n,$$

and from (3) and (8) in the interval (7)

$$|P(h)| \leqq M_0 + \frac{(b-a)^{n+1}}{(n+1)!}\cdot M_{n+1} = K. \tag{10}$$

By applying Markoff's theorem to the polynomial $P(h)$ we find

$$|P'(h)| \leqq \frac{2n^2}{b-a}\cdot K$$

and, since $h=0$ belongs to the interval (7), we have

$$|P'(0)| = |f'(x)| \leqq \frac{2n^2}{b-a}\cdot K.$$

3. When $f(x)$ is a polynomial, Theorem 1 obviously reduces to the theorem of A. Markoff. One may state Theorem 1 briefly by saying that when $f(x)$ and $f^{(n+1)}(x)$ are bounded in (a, b) then $f'(x)$ has the same property. By repetition one obtains a bound for all intermediate derivatives

* Professor Hille pointed out the advantage in using this form for the remainder term.

$$f^{(i)}(x) \qquad\qquad (i = 1, 2, \cdots, n).$$

A better bound is obtained however by applying the preceding extension principle directly to a more general theorem by W. Markoff:*

Let $f_n(x)$ be a polynomial of degree n and M_0 the maximum of $|f_n(x)|$ in the interval (a, b). Then one has in the same interval

$$(11)\qquad \left| f_n^{(i)}(x) \right| \leqq K(i, n)\cdot\frac{M_0}{(b-a)^i},$$

where

$$(12)\qquad K(i, n) = \frac{2^i\cdot n^2(n^2-1)\cdots(n^2-(i-1)^2)}{1\cdot 3\cdot 5\cdots(2i-1)} = \frac{n}{n+i}\cdot 2^{2i}\cdot i! C_{n+i,2i}.$$

The equality sign can only hold for the Tschebyschef polynomials (2).

When this result is applied to the derivatives of the polynomial (9) in the interval (7) we obtain:

THEOREM 2. *Let $f(x)$ possess an $(n+1)$st derivative $f^{(n+1)}(x)$ such that in the interval (a, b)*

$$|f(x)| \leqq M_0, \qquad |f^{(n+1)}(x)| \leqq M_{n+1}.$$

Then all intermediate derivatives of $f(x)$ are bounded in the same interval by

$$(13)\qquad |f^{(i)}(x)| \leqq \frac{n}{n+i}\cdot 2^{2i}\cdot i!\cdot C_{n+i,2i}\frac{1}{(b-a)^i}\cdot\left(M_0 + M_{n+1}\cdot\frac{(b-a)^{n+1}}{(n+1)!}\right).$$

Let us observe that the results of W. Markoff were proved only under the assumption that the polynomial $f(x)$ has real coefficients. One may however easily extend the results to complex coefficients and hence Theorem 2 to complex valued functions.

If one introduces the notation

$$(14)\qquad |f^{(i)}(x)| \leqq M_i, \qquad \left|\frac{f^{(i)}(x)}{i!}\cdot(b-a)^i\right| \leqq \overline{M}_i,$$

then

$$(15)\qquad \overline{M}_i = \frac{M_i}{i!}(b-a)^i$$

and the inequality (13) may be written in the simpler form

$$(16)\qquad \overline{M}_i \leqq \frac{n}{n+i}\cdot 2^{2i}\cdot C_{n+i,2i}\cdot(\overline{M}_0 + \overline{M}_{n+1}).$$

* W. Markoff, *Über Polynome die in einem gegebenen Intervalle möglichst wenig von Null abweichen*, Mathematische Annalen, vol. 77 (1916), pp. 213–258.

This relation shows that the constants M_i or $\overline{M}_i$ for a real, differentiable function satisfy certain restricting conditions. It suggests the very interesting problem of determining the necessary and sufficient condition in order that a series of positive numbers

$$M_0, M_1, \cdots$$

be the maxima of the derivatives of a function $f(x)$ in an interval.

4. Let us next turn to the case of analytic functions. Let us suppose that $f(z)$ is analytic and regular in a certain domain D in the complex plane. Furthermore, $f^{(n+1)}(z)$ is bounded in the same domain. Let D be bounded by a Jordan curve C such that one may draw through each point in D a chord of length $d>0$ entirely contained in D. When Theorem 1 is applied to the chords of D, one finds that in D

$$|f'(z)| \leqq \frac{2n^2}{d}\cdot\left(M_0 + \frac{d^{n+1}}{(n+1)!}\cdot M_{n+1}\right).$$

This remark gives extensions of results obtained by Jackson,* Sewell,† and others. For the higher derivatives one finds corresponding to (13)

$$|f^{(i)}(z)| \leqq \frac{1}{d^i}\cdot K(i, n)\cdot K,$$

where $K(i, n)$ is given by (12).

One may however obtain considerably better results through another procedure. For polynomials in the complex plane we have the following theorem of S. Bernstein:‡

Let $f_n(z)$ be a polynomial of degree n and let M_0 denote the maximum of $f_n(z)$ on a circle with radius R. Then we have on the same circle

$$|f_n'(z)| \leqq n\cdot\frac{M_0}{R}\cdot \tag{17}$$

The equality sign holds only for

$$f_n(z) = M_0\cdot e^{i\theta}\cdot\left(\frac{z-a}{R}\right)^n. \tag{18}$$

* D. Jackson, *On the application of Markoff's theorem to problems of approximation in the complex domain*, Bulletin of the American Mathematical Society, vol. 37 (1931), pp. 883–890.

† W. E. Sewell, *Generalized derivatives and approximation by polynomials*, these Transactions, vol. 41 (1937), pp. 84–123. This paper gives further references particularly to Szegö and Montel.

‡ S. Bernstein, *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle*, Paris, 1926. See also M. Riesz, *Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 23 (1914), pp. 354–368.

By means of the same extension principle which we used in the preceding this theorem may be extended to arbitrary analytic functions. Let $f(z)$ be a function analytic on the circle C. It is not necessary to assume that $f(z)$ is regular in the interior of C. The results hold even when $f(z)$ is a branch of an analytic function not returning to its original value by a circuit of C.

In all cases there exists a Taylor expansion

$$f(z+h) = f(z) + \frac{h}{1!}f'(z) + \cdots + \frac{h^n}{n!}f^{(n)}(z) + R_n, \tag{19}$$

where

$$R_n = \frac{(-1)^n}{n!}\int_z^{z+h}(t-z-h)^n f^{(n+1)}(t)dt. \tag{20}$$

Here z and $z+h$ are points on C and the path of integration is taken along C in some fixed direction. Furthermore $f(z+h)$ is the value of $f(z)$ determined by the path. Let us now suppose that for the chosen branch of $f(z)$ we have

$$|f(z)| \leqq M_0, \qquad |f^{(n+1)}(z)| \leqq M_{n+1}$$

for the points of C. For the remainder term R_n in (20) one finds the estimate

$$|R_n| \leqq \frac{2}{n!}(2R)^{n+1}\cdot M_{n+1}\cdot \int_0^{\pi/2}\cos^n \phi d\phi.$$

The last integral tends to zero with increasing n. Let us use however only the rough estimate

$$|R_n| < \frac{2}{n!}(2R)^{n+1}\cdot M_{n+1}. \tag{21}$$

Now let z be a fixed point on C. Since $z+h$ is also located on C the point h describes another circle C' with the same radius. Let us write again

$$P(h) = f(z) + \frac{h}{1!}f'(z) + \cdots + \frac{h^n}{n!}f^{(n)}(z). \tag{22}$$

According to (19) and (21) we have for the points on the circle C'

$$|P(h)| \leqq M_0 + \frac{2}{n!}(2R)^{n+1}\cdot M_{n+1} = Q. \tag{23}$$

By applying the theorem of Bernstein, we obtain

$$|P'(h)| \leqq \frac{n}{R}\cdot Q$$

and, since $h=0$ is located on C',

$$| P'(0) | = | f'(z) | \leqq \frac{n}{R} \cdot Q.$$

By specialization to the unit circle this theorem may be stated as follows:

THEOREM 3. *Let $f(z)$ be a function which is analytic on the unit circle C_1. If then*

$$| f(z) | \leqq M_0, \qquad | f^{(n+1)}(z) | \leqq M_{n+1}$$

on C_1, then

$$| f'(z) | \leqq n \left(M_0 + \frac{2^{n+1}}{n!} M_{n+1} \right)$$

on the unit circle.

The generalization of this theorem to higher derivatives implies the following extension of the theorem of Bernstein:

THEOREM 4. *Let $f_n(z)$ be a polynomial of degree n and let M_0 denote its maximum on a circle with radius R. Then one has on the same circle*

$$| f_n^{(i)}(z) | \leqq n(n-1) \cdots (n-i+1) \cdot \frac{M_0}{R^i} \cdot$$

The equality sign can hold only for the polynomials (18).

To prove Theorem 4 it is only necessary to apply the theorem of Bernstein i times.

When Theorem 4 is applied to the polynomial $P(h)$ in (22) we obtain

$$| P^{(i)}(h) | \leqq \frac{n!}{(n-i)!} \cdot \frac{Q}{R^i},$$

where Q is defined by (23). For $h=0$ we find the desired result

$$| f^{(i)}(z) | \leqq \frac{n!}{(n-i)!} \cdot \frac{1}{R^i} \left(M_0 + \frac{2^{n+1}}{n!} M_{n+1} \right).$$

THEOREM 5. *Let $f(z)$ be analytic on the unit circle and*

$$| f(z) | \leqq M_0, \qquad | f^{(n+1)}(z) | \leqq M_{n+1}$$

on the circle. One then also has

$$| f^{(i)}(z) | \leqq \frac{n!}{(n-i)!} \left(M_0 + \frac{2^{n+1}}{n!} M_{n+1} \right), \qquad (i = 1, 2, \cdots, n).$$

YALE UNIVERSITY,
NEW HAVEN, CONN.

COMPARISON OF PRODUCTS OF METHODS OF SUMMABILITY*

BY
RALPH PALMER AGNEW

1. **Introduction.** A sequence s_n of complex numbers (or complex-valued functions) is called summable to L by the method of summability

$$(A) \qquad S_n = \sum_{k=1}^{\infty} a_{nk}s_k,$$

determined by the matrix $A \equiv (a_{nk})$ of real or complex constants, if the transform S_n exists and $\lim_{n\to\infty} S_n = L$. The matrix A (and method of summability A) is called row-finite if for each n, $a_{nk}=0$ for all sufficiently great k; and is called triangular if $a_{nk}=0$ for $k>n$. The method A is regular if $s_n \to L$ implies $S_n \to L$. Necessary and sufficient conditions that A be regular are, by the Silverman-Toeplitz theorem,

$$(1.1) \qquad \sum_{k=1}^{\infty} | a_{nk} | < M, \qquad M = \text{constant},$$

$$(1.2) \qquad \text{for each } k, \quad \lim_{n\to\infty} a_{nk} = 0,$$

$$(1.3) \qquad \lim_{n\to\infty} \sum_{k=1}^{\infty} a_{nk} = 1.$$

The set of sequences summable A is called the convergence field of A.

Let

$$(B) \qquad T_n = \sum_{k=1}^{\infty} b_{nk}s_k$$

denote a second method of summability.

In case each sequence summable B is summable A to the same value, A is said to include B and we write $A \supset B$. In case $A \supset B$ and $B \supset A$, A and B are called equivalent and we write $A \sim B$. In case the equality $L_A = L_B$ holds for each sequence summable A to L_A and summable B to L_B, the methods A and B are called mutually consistent (or consistent).

In terms of A and B it is possible to define two "products," each of which is a new method of summability. The *iteration product*, ordinarily denoted by

* Presented to the Society, February 20, 1937; received by the editors March 27, 1937.

AB, is the method which associates with a given sequence the A transform of its B transform, that is,

$$(AB) \qquad U_n = \sum_{p=1}^{\infty} a_{np}T_p = \sum_{p=1}^{\infty}\sum_{k=1}^{\infty} a_{np}b_{pk}s_k.$$

Thus s_n is summable AB to L if $\lim U_n = L$. The *composition product* is also at times denoted by AB; it is the method whose matrix is the product AB (which we denote by $A \cdot B$) of the matrices A and B. Thus we write

$$(A \cdot B) \qquad V_n = \sum_{k=1}^{\infty}\sum_{p=1}^{\infty} a_{np}b_{pk}s_k,$$

and s_n is summable $A \cdot B$ to L if $V_n \to L$.

We observe that U_n and V_n are, if they exist, respectively the "sum by rows" and the "sum by columns" of the double series

$$(1.4) \qquad \begin{aligned} &a_{n1}b_{11}s_1 + a_{n1}b_{12}s_2 + a_{n1}b_{13}s_3 + \cdots \\ &+ a_{n2}b_{21}s_1 + a_{n2}b_{22}s_2 + a_{n2}b_{23}s_3 + \cdots \\ &+ a_{n3}b_{31}s_1 + a_{n3}b_{32}s_2 + a_{n3}b_{33}s_3 + \cdots \\ &+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots . \end{aligned}$$

If A and B are regular and s_n is bounded, the series (1.4) converges absolutely and $U_n = V_n$; but without these restrictions it is not so obvious that $U_n = V_n$. There is in fact the possibility that AB and $A \cdot B$ may fail to be equivalent or even consistent.

It is the main object of this paper to compare pairs selected from the four transformations A, B, AB, and $A \cdot B$, considering in each case questions of inclusion, equivalence, and consistency. It appears that unless either or both of the matrices (a_{nk}) and (b_{nk}) are assumed to belong to restricted types, the results obtained are largely negative. These negative results are established by examples. Several examples are explicitly given, each for two reasons. In the first place each example, consisting of two regular methods of summability satisfying prescribed conditions and a sequence, can be manufactured only after considerable experimentation. In the second place the examples are largely of such obviously pathological character that they leave hope of obtaining positive theorems involving matrices of restricted types. Some such theorems are given in this paper, particularly in §11. It is doubtless true that more (and better) theorems of this kind will appear in the future.

In §12 we compare AB with $A'B'$ and $A \cdot B$ with $A' \cdot B'$ where the pair A, A' and the pair B, B' represent closely related methods of summability.

In §13 we deal briefly with multiple products and in §14 with kernel transformations.

2. **Comparison of methods A and B.** It is well known that two regular row-finite methods A and B with $a_{nk}\geqq 0$, $b_{nk}\geqq 0$ may be such that all of the relations $A\supset B$, $B\supset A$, $A\sim B$ are false and in fact such that A and B are inconsistent.*

3. **Comparison of methods A and AB.** It follows from §2 that we can choose regular row-finite methods A and B, with $a_{nk}\geqq 0$, $b_{nk}\geqq 0$, and a sequence s_n summable A to L_A and summable B to $L_B\neq L_A$. It is easy to see that regularity of A implies that s_n is summable AB to L_B. Thus A and AB may be inconsistent, and there is no hope of showing that $A\supset AB$, $AB\supset A$ or $A\sim AB$.

4. **Comparison of B and AB.** Elementary examples show that $B\supset AB$ may be false, and hence that B and AB need not be equivalent, even when A and B are assumed to be regular, row-finite and $a_{nk}\geqq 0$, $b_{nk}\geqq 0$.

However if A is regular and s_n is summable B to L, then $T_n\to L$ and regularity of A imply $U_n\to L$ so that s_n is summable AB to L. It follows that if A is regular, then $AB\supset B$. This implies that B and AB must be consistent.

5. **Comparison of A and $A\cdot B$.** When A and B are determined as in §3, the series (1.4) from which U_n and V_n are computed reduces to a finite sum, and obviously $U_n=V_n$; hence in this case $A\cdot B\sim AB$. It follows from §3 that A and $A\cdot B$ may be inconsistent, and there is no hope of showing that $A\supset A\cdot B$, $A\cdot B\supset A$, or $A\sim A\cdot B$.

6. **Comparison of B and $A\cdot B$.** Elementary examples of row-finite transformations show that $B\supset A\cdot B$ may be false and hence that B and $A\cdot B$ need not be equivalent.

One might expect to be able to show that if A and B are regular, then $A\cdot B\supset B$. But this is impossible. The author† has given an example of transformations A and B (having some significant properties in addition to regularity) and a sequence s_n which is summable B but non-summable $A\cdot B$.

In this paper we go further and prove in §9 the following theorem:

THEOREM 6.1. *There exist regular transformations A and B with $a_{nk}\geqq 0$, $b_{nk}\geqq 0$ for which B and $A\cdot B$ are inconsistent.*

7. **Comparison of AB and $A\cdot B$.** It is not true that regularity of A and B

* Questions involving inconsistency of transformations are discussed in detail in Agnew, *On ranges of inconsistency of regular transformations and allied topics*, Annals of Mathematics, vol. 32 (1931), pp. 715–722.

† Agnew, *Products of methods of summability*, Bulletin of the American Mathematical Society, vol. 42 (1936), pp. 547–549.

implies either $AB \supset A \cdot B$ or $A \cdot B \supset AB$. In fact we prove in §9 the following theorem:

THEOREM 7.1. *There exist regular transformations A and B with $a_{nk} \geqq 0$, $b_{nk} \geqq 0$, for which AB and $A \cdot B$ are inconsistent.*

This theorem evidences the necessity of noticing a distinction between the iteration product AB and the composition product $A \cdot B$.

The transformation A of Theorem 7.1 cannot be row-finite. For if A is row-finite, then all of the terms lying below some row of the double series (1.4) from which U_n and V_n are computed vanish, and existence of U_n implies existence of V_n and the equality $V_n = U_n$. Thus we have the theorem:

THEOREM 7.2. *If A is row-finite, then $A \cdot B \supset AB$.*

This implies that if A is row-finite, then AB and $A \cdot B$ must be consistent. It is however impossible to go further and prove that AB and $A \cdot B$ must be equivalent. We prove the following theorem:

THEOREM 7.3. *There exist regular transformations A and B with A row-finite, $a_{nk} \geqq 0$, $b_{nk} \geqq 0$, and a sequence s_k such that s_k is summable $A \cdot B$ but non-summable AB.*

Let $p_1, p_2, p_3, \cdots$ denote in order the primes $2, 3, 5, \cdots$. For each $n = 1, 2, 3, \cdots$, let $a_{n,p_n} = a_{n,p_n^2} = 1/2$ and let $a_{nk} = 0$ otherwise. If n is neither a prime nor a square of a prime, let $b_{nn} = 1$, and $b_{nk} = 0$ otherwise. For each $n = 1, 2, \cdots$ let $b_{p_n,k} = 0$ when $k \neq p_n, p_n^3, p_n^5, \cdots$, and let $b_{p_n,k} = 2^{-\alpha}$ when k is of the form $p_n^{2\alpha-1}$. Let $b_{p_n^2,k} = 0$ when $k \neq p_n^2, p_n^4, p_n^6, \cdots$, and let $b_{p_n^2,k} = 2^{-\alpha}$ when k is of the form $p_n^{2\alpha}$. These matrices a_{nk} and b_{nk} define methods A and B of summability satisfying the hypotheses of the theorem. Observe that $a_{nk} = b_{nk} = 0$ when $n > k$. Let the sequence s_k be defined by the formulas: $s_k = 2^{\alpha+1}/\alpha$ when k is of the form $p_n^{2\alpha-1}$; $s_k = -2^{\alpha+1}/\alpha$ when k is of the form $p_n^{2\alpha}$; and $s_k = 0$ otherwise.

It can be shown that for this example the double series (1.4) from which U_n and V_n are computed becomes (after omission of rows and columns of zeros)

$$\begin{aligned} &1 + 0 + \tfrac{1}{2} + 0 + \tfrac{1}{3} + 0 + \tfrac{1}{4} + \cdots \\ &+ 0 - 1 + 0 - \tfrac{1}{2} + 0 - \tfrac{1}{3} + 0 - \cdots . \end{aligned} \tag{7.31}$$

It is apparent that, for each n, the sum by columns of this series is 0, that is, $V_n = 0$; and that the sum by rows does not exist, that is, U_n does not exist. Thus the sequence of s_n of the example is summable $A \cdot B$ to 0 and is non-summable AB. This proves Theorem 7.3.

The transformation B of Theorem 7.3 cannot be row-finite. For if both

A and B are row-finite, then $U_n = V_n$ for every n and equivalence of AB and $A \cdot B$ follows.

In spite of the fact that the transformations AB and $A \cdot B$ of Theorem 7.1 need not be consistent, there is a large class of sequences (including all bounded sequences and all unilaterally bounded real sequences) over which they must be equivalent.

We shall say that a sequence s_n lies in an angle less than π in the complex plane if there exist a point z_0, an angle θ_0, and a positive angle $\phi < \pi/2$ such that for each n

$$s_k = z_0 + \rho_k e^{i(\theta_0 + \theta_k)} \tag{7.32}$$

where $\rho_k \geqq 0$ and $|\theta_k| \leqq \phi$.

THEOREM 7.4. *If A and B are regular transformations with*

$$a_{nk} \geqq 0, \qquad b_{nk} \geqq 0, \qquad n, k = 1, 2, \cdots, \tag{7.41}$$

then each sequence s_n which lies in an angle less than π in the complex plane and which is summable to L by one of the methods AB and $A \cdot B$ is also summable to L by the other one.

The gist of this theorem is that for regular transformations A, B satisfying (7.41), the two transformations AB and $A \cdot B$ are equivalent in so far as application to sequences lying in an angle less than π is concerned.

To prove the theorem, suppose first that s_k is a sequence given by (7.32) for which V_n exists. Then

$$V_n = \sum_{k=1}^{\infty} \sum_{p=1}^{\infty} a_{np} b_{pk} [z_0 + \rho_k e^{i\theta_0} e^{i\theta_k}], \tag{7.42}$$

and, where ξ_k and η_k are the real and imaginary parts of $\rho_k e^{i\theta_k}$,

$$V_n = z_0 \sum_{k=1}^{\infty} \sum_{p=1}^{\infty} a_{np} b_{pk} + e^{i\theta_0} \sum_{k=1}^{\infty} \sum_{p=1}^{\infty} a_{np} b_{pk} [\xi_k + i\eta_k]. \tag{7.43}$$

Since $a_{nk} \geqq 0$, $b_{nk} \geqq 0$, $\xi_k \geqq 0$, and η_k is real, this implies convergence of

$$\sum_{k=1}^{\infty} \sum_{p=1}^{\infty} a_{np} b_{pk} |\xi_k|. \tag{7.44}$$

But $|\eta_k| \leqq \xi_k \tan \phi$ so that (7.44) converges when $|\xi_k|$ is replaced by $|\eta_k|$. It now follows easily that both series in (7.43) converge absolutely and hence that the series in (7.42) converges absolutely. Therefore

$$U_n = \sum_{p=1}^{\infty} \sum_{k=1}^{\infty} a_{np} b_{pk} [z_0 + \rho_k e^{i\theta_0} e^{i\theta_k}] \tag{7.45}$$

exists and $U_n = V_n$; hence $V_n \to L$ implies also $U_n \to L$. We can show similarly that $U_n \to L$ implies also $V_n \to L$, and Theorem 7.4 is proved.

8. **A double series.** In the next section we use the double series

$$
(8.1)\qquad
\begin{array}{l}
1 - \tfrac{1}{2} + 0 - \tfrac{1}{4} + 0 - \tfrac{1}{6} + 0 - 0 + 0 - 0 \ + \cdots \\
+\, 0 - 0 + \tfrac{1}{3} - 0 + 0 - 0 + 0 - \tfrac{1}{8} + 0 - \tfrac{1}{10} + \cdots \\
+\, 0 - 0 + 0 - 0 + \tfrac{1}{5} - 0 + 0 - 0 + 0 - 0 \ + \cdots \\
+\, 0 - 0 + 0 - 0 + 0 - 0 + \tfrac{1}{7} - 0 + 0 - 0 \ + \cdots \\
+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{array}
$$

whose terms u_{nk} may be defined as follows: for each odd k

$$
(8.2)\qquad
\begin{aligned}
u_{nk} &= 1/k \ \text{ if } n = (k+1)/2, \\
&= 0 \ \text{ otherwise;}
\end{aligned}
$$

and for each even k

$$
(8.3)\qquad
\begin{aligned}
u_{nk} &= -\,1/k \ \text{ if } n = n_k, \\
&= 0 \ \text{ otherwise,}
\end{aligned}
$$

where n_k is the smallest n for which

$$
(8.4)\qquad u_{n1} + u_{n2} + u_{n3} + \cdots + u_{n,k-1} - 1/k \geqq 0.
$$

The harmonic series $\sum 1/k$ being divergent, it is easy to see that for each n, the infinite series $u_{n1}+u_{n2}+\cdots$ converges to 0. Hence the double series (8.1) converges by rows to 0 and by columns to $\log 2 = 1-1/2+1/3-1/4+\cdots$. Moreover it can be shown by arithmetic methods that for each n, $u_{nk}=0$ for all sufficiently great k; that is, each row of (8.1) contains only a finite number of non-vanishing terms.

9. **Proof of Theorems 6.1 and 7.1.** We can now prove the following theorem of which Theorems 6.1 and 7.1 are obvious corollaries:

THEOREM 9.1. *There exists a pair of regular transformations A and B with B row-finite,*

$$
(9.11)\qquad a_{nk} \geqq 0, \qquad b_{nk} \geqq 0, \qquad n, k = 1, 2, \cdots,
$$

$$
(9.12)\qquad \sum_{k=1}^{\infty} a_{nk} = 1, \qquad \sum_{k=1}^{\infty} b_{nk} = 1, \qquad n = 1, 2, \cdots,
$$

and a sequence s_k such that

$$
(9.13)\qquad T_n = \sum_{k=1}^{\infty} b_{nk} s_k = 0, \qquad n = 1, 2, \cdots,
$$

$$U_n = \sum_{p=1}^{\infty} a_{np}T_p = \sum_{p=1}^{\infty}\sum_{k=1}^{\infty} a_{np}b_{pk}s_k = 0, \qquad n = 1, 2, \cdots, \tag{9.14}$$

and

$$V_n = \sum_{k=1}^{\infty}\sum_{p=1}^{\infty} a_{np}b_{pk}s_k = 1, \qquad n = 1, 2, \cdots. \tag{9.15}$$

Let the positive integers 1, 2, 3, $\cdots$ be displayed as a double sequence h_{nk} so that $h_{11}=1$, $h_{21}=2$, $h_{12}=3$, $h_{31}=4$, $h_{22}=5$, $h_{13}=6$, $h_{14}=7$, etc. For each $n=1, 2, 3, \cdots$, let a_{nk} be defined for $k=1, 2, 3, \cdots$ by the formula

$$\begin{aligned} a_{nk} &= 0, \qquad && k \neq h_{n1}, h_{n2}, h_{n3}, \cdots, \\ &= 2^{-r}, && k = h_{nr},\ r = 1, 2, \cdots. \end{aligned} \tag{9.16}$$

This matrix a_{nk} determines a regular method A of summability with $a_{nk} \geqq 0$ and $\sum_{k=1}^{\infty} a_{nk}=1$ for each n.

The double series (1.4) from which U_n and V_n are computed takes (after removal of rows of zeros) the form

$$\begin{aligned} &\tfrac{1}{2}b_{h_{n1},1}s_1 + \tfrac{1}{2}b_{h_{n1},2}s_2 + \tfrac{1}{2}b_{h_{n1},3}s_3 + \cdots \\ &+ \tfrac{1}{4}b_{h_{n2},1}s_1 + \tfrac{1}{4}b_{h_{n2},2}s_2 + \tfrac{1}{4}b_{h_{n2},3}s_3 + \cdots \\ &+ \tfrac{1}{8}b_{h_{n3},1}s_1 + \tfrac{1}{8}b_{h_{n3},2}s_2 + \tfrac{1}{8}b_{h_{n3},3}s_3 + \cdots \\ &+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots . \end{aligned} \tag{9.17}$$

Let, for each n and r,

$$b_{h_{n,r},k} = 0, \qquad k \neq h_{n,r}, h_{n,r+1}, \cdots. \tag{9.18}$$

Then the double series (9.17) takes (after removal of columns of zeros) the form

$$\begin{aligned} &\tfrac{1}{2}b_{h_{n1},h_{n1}}s_{h_{n1}} + \tfrac{1}{2}b_{h_{n1},h_{n2}}s_{h_{n2}} + \tfrac{1}{2}b_{h_{n1},h_{n3}}s_{h_{n3}} + \cdots \\ +\ &\quad 0 \quad + \tfrac{1}{4}b_{h_{n2},h_{n2}}s_{h_{n2}} + \tfrac{1}{4}b_{h_{n2},h_{n3}}s_{h_{n3}} + \cdots \\ +\ &\quad 0 \quad + \quad 0 \quad + \tfrac{1}{8}b_{h_{n3},h_{n3}}s_{h_{n3}} + \cdots \\ +\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots . \end{aligned} \tag{9.19}$$

We observe that if $n' \neq n''$, then the "variables" $b_{\alpha\beta}$ and s_γ appearing in (9.19), when $n=n'$, are distinct from those appearing when $n=n''$. For each $n=1, 2, 3, \cdots$ let the elements $b_{\alpha\beta}$ and s_γ appearing in (9.19) be determined so that the terms of the two series (9.19) and (8.1) in corresponding positions will be equal, and the non-vanishing elements of the sequence

$$b_{h_{n,r}h_{n,r}}, b_{h_{n,r}h_{n,r+1}}, b_{h_{n,r}h_{n,r+2}}, \cdots$$

will be equal in order to

$$1/2d,\ 1/2^2d, \cdots, 1/2^5d$$

where $d=1/2+1/2^2+\cdots+1/2^5$.

This gives a complete and unique determination of the elements of the matrix $B\equiv(b_{nk})$. It is clear that B satisfies the hypotheses of the theorem, regularity being implied by the conditions

$$b_{nk} \geqq 0, \qquad \sum_{k=1}^{\infty} b_{nk} = 1,$$

and the fact that $b_{nk}=0$ when $n>k$.

It follows from identity of (9.19) and (8.1) and the fact that each row of (8.1) converges to 0, that $T_p=0$ for each $p=1, 2, \cdots$; and since (8.1) and hence (9.19) converge by rows to 0 and by columns to log 2 we have $U_n=0$ and $V_n=\log 2$. If finally we divide each s_n determined above by log 2, then T_p, U_n, and V_n will be divided by log 2 and we obtain (9.13), (9.14), and (9.15).

10. **Remarks on Theorem** 9.1. The author has been unable to find an example less recondite than the one just given to prove Theorem 9.1. In case the requirements $a_{nk}\geqq 0$, $b_{nk}\geqq 0$ are removed, we can give simpler examples. For definiteness, and convenience of reference we state the following theorem:

THEOREM 10.1. *If r is a complex number with $0<|r|<1$, then the methods*

$$(A) \qquad S_n = s_n + r^n s_{n+1} + r^{n+1}s_{n+2} + \cdots,$$

$$(B) \qquad T_n = [1/(1-r)]s_{n-1} + [-r/(1-r)]s_n$$

are regular while AB and $A\cdot B$ are inconsistent. The sequence $s_k=(1-r)r^{-k}$ is summable AB to 0 *and $A\cdot B$ to* 1.

Verification is straightforward and left to the reader. We note that if $0<r<1$, the elements a_{nk} are all $\geqq 0$ but some elements b_{nk} are <0; while if $-1<r<0$, all elements b_{nk} are $\geqq 0$ but some elements a_{nk} are <0. If r is not real, the conditions $a_{nk}\geqq 0$, $b_{nk}\geqq 0$ both fail.

The method A of Theorem 10.1 is, for each admissible r, equivalent to convergence. For on one hand A is regular. On the other hand if s_k is summable A to L so that

$$(10.11) \qquad S_n = s_n + r^n s_{n+1} + r^{n+1}s_{n+2} + \cdots$$

exists and $S_n\to L$, then convergence of the series in (10.11) implies that

$$(10.12) \qquad \lim_{\alpha\to\infty} (r^\alpha s_{\alpha+1} + r^{\alpha+1}s_{\alpha+2} + \cdots) = 0,$$

and hence that

$$(10.13) \qquad \lim s_n = \lim S_n = L.$$

The method B is not only regular but also has several other features at times desirable in methods of summability. The permissibility of removal or adjunction of elements at the beginning of a sequence is such a feature.

These remarks make it appear likely that significant theorems giving conditions sufficient for consistency of AB and $A\cdot B$ (or for $AB\supset A\cdot B$, or for $A\cdot B\supset AB$, or for $AB\sim A\cdot B$) will involve classes of methods defined by matrices of more or less restricted types rather than involve classes of methods having various ones of the numerous "desirable" properties of methods of summability.

The following theorem indicates the possibility of obtaining constructive theorems involving AB and $A\cdot B$, and is of interest in connection with Theorem 10.1:

THEOREM 10.2. *If A and B are regular transformations with $a_{nk}\geqq 0$, $b_{nk}\geqq 0$ and if B is of the form*

$$T_n=\sum_{k=n-\alpha}^{n+\beta} b_{nk}s_k,$$

where α and β are non-negative integers, then $AB\supset A\cdot B$.

In interpreting B, we agree that $s_k=0$ when $k<1$, and that $b_{pk}=0$ when $k<p-\alpha$ and when $k>p+\beta$. Assuming s_k to be a sequence for which

$$V_n=\sum_{k=1}^{\infty}\sum_{p=1}^{\infty} a_{np}b_{pk}s_k\equiv\sum_{k=1}^{\infty}\sum_{p=k-\beta}^{k+\alpha} a_{np}b_{pk}s_k \tag{10.21}$$

exists, we have for fixed n

$$\begin{aligned} V_n&=\lim_{Q\to\infty}\sum_{k=1}^{Q}\sum_{p=1}^{\infty} a_{np}b_{pk}s_k=\lim_{Q\to\infty}\sum_{p=1}^{\infty} a_{np}\sum_{k=1}^{Q} b_{pk}s_k\\ &=\lim_{Q\to\infty}\left\{\sum_{p=1}^{Q-\beta} a_{np}\sum_{k=1}^{Q} b_{pk}s_k+\sum_{p=Q-\beta+1}^{Q+\alpha} a_{np}\sum_{k=1}^{Q} b_{pk}s_k\right\}, \end{aligned} \tag{10.22}$$

and hence

$$V_n=\lim_{Q\to\infty}\left\{\sum_{p=1}^{Q-\beta} a_{np}\sum_{k=1}^{\infty} b_{pk}s_k+\sum_{p=Q-\beta+1}^{Q+\alpha}\ \sum_{k=Q-\alpha-\beta+1}^{Q} a_{np}b_{pk}s_k\right\}. \tag{10.23}$$

The operations under the limit sign are justified by vanishing of elements b_{nk}. Now convergence of the first series in (10.21) implies that

$$\Delta_k\equiv\sum_{p=1}^{\infty} a_{np}b_{pk}\,|\,s_k\,|\to 0$$

as $k\to\infty$. But since $a_{nk}\geqq 0$ and $b_{nk}\geqq 0$, $0\leqq a_{np}b_{pk}|s_k|\leqq \Delta_k$ for each fixed p.

Hence

$$\left|\sum_{p=Q-\beta+1}^{Q+\alpha}\ \sum_{k=Q-\alpha-\beta+1}^{Q} a_{np}b_{pk}s_k\right| \leqq \sum_{p=Q-\beta+1}^{Q+\alpha}\ \sum_{k=Q-\alpha-\beta+1}^{Q} \Delta_k = (\alpha+\beta)\sum_{k=Q-\alpha-\beta+1}^{Q}\Delta_k \to 0$$

as $Q\to\infty$. This fact and (10.23) imply that

$$U_n = \sum_{p=1}^{\infty} a_{np}\sum_{k=1}^{\infty} b_{pk}s_k \equiv \sum_{p=1}^{\infty} a_{np}\sum_{k=p-\alpha}^{p+\beta} b_{pk}s_k$$

exists and $U_n = V_n$. This argument shows that $AB\supset A\cdot B$, and Theorem 10.2 is proved.

Examples show it is impossible to modify the argument to prove $A\cdot B\supset AB$. For one such example, it suffices to put, for each $n=1, 2, \cdots$,

$$\begin{aligned} a_{nk} &= 0, & k &\neq n^2, n^4, n^8, \cdots, \\ &= 2^{-\sigma}, & k &= n^{2^\sigma}, & \sigma &= 1, 2, \cdots, \end{aligned} \tag{10.4}$$

and for each $n=2, 3, \cdots$,

$$\begin{aligned} b_{nk} &= 0, & k &\neq n-1, n, \\ &= 1/2, & k &= n-1, n, \end{aligned} \tag{10.5}$$

while $b_{11}=1$ and $b_{1k}=0$ for $k>1$. The sequence defined by $s_n=(-1)^n\log n$ is summable AB to 0 and is non-summable $A\cdot B$. We note in passing that (10.5) defines a regular Nörlund method of summability which is included by the arithmetic mean method.

11. **The arithmetic mean and generalizations.** Let $p_1, p_2, p_3, \cdots$ be a sequence of constants with

$$P_n = p_1+p_2+\cdots+p_n \neq 0, \qquad n=1, 2, \cdots. \tag{11.01}$$

Let P denote the method of summability associated with the transformation

$$T_n = (p_1s_1+p_2s_2+\cdots+p_ns_n)/P_n. \tag{P}$$

The transformations P differ from the more familiar Nörlund methods in order of distribution of the "weights" p_k; but share with Nörlund methods the property of reducing to the important arithmetic mean method C_1 (or M) when $p_k=1$ for each k. The theorems of this section therefore give facts involving C_1.

THEOREM 11.1. *If A and P (regular or not) are methods of summability with $a_{nk}\geqq 0$, $P_n>0$, then $AP\supset A\cdot P$.*

The AP and $A\cdot P$ transforms U_n and V_n of a sequence s_k are (if they exist) determined as the sum by rows and the sum by columns of the double series obtained by removing parentheses from the series

$$
\begin{aligned}
&P_1^{-1}a_{n1}(p_1s_1 + 0 + 0 + 0 + \cdots) \\
&+ P_2^{-1}a_{n2}(p_1s_1 + p_2s_2 + 0 + 0 + \cdots) \\
&+ P_3^{-1}a_{n3}(p_1s_1 + p_2s_2 + p_3s_3 + 0 + \cdots) \\
&+ \cdots\cdots\cdots\cdots\cdots\cdots .
\end{aligned} \tag{11.10}
$$

We show that existence of V_n implies existence of U_n and the equality $U_n = V_n$. This follows, on introducing obvious notation and interchanging rows and columns, from the following lemma:

LEMMA 11.2. *If $\theta_n \geqq 0$ for $n=1, 2, \cdots$, and the series obtained by removing parentheses from the series*

$$
\begin{aligned}
&\sigma_1(\theta_1 + \theta_2 + \theta_3 + \cdots) \\
&+ \sigma_2(\qquad \theta_2 + \theta_3 + \cdots) \\
&+ \sigma_3(\qquad\qquad \theta_3 + \cdots) \\
&+ \cdots\cdots\cdots\cdots
\end{aligned} \tag{11.20}
$$

converges by rows to Λ, then it also converges by columns to Λ.

To prove the lemma, let

$$R_n = \theta_n + \theta_{n+1} + \theta_{n+2} + \cdots . \tag{11.21}$$

If $R_n = 0$ for some n, then $\theta_n = 0$ for all sufficiently great n, and the conclusion of the lemma obviously holds. Hence we may assume $R_n \neq 0$ for all n. Let

$$\omega_n = \sigma_n(\theta_n + \theta_{n+1} + \theta_{n+2} + \cdots) = \sigma_n R_n. \tag{11.22}$$

Then

$$\sum \omega_n = \Lambda, \qquad \sigma_n = \omega_n / R_n. \tag{11.23}$$

The sum of n columns of the series (11.20) is

$$K_n = \sigma_1(R_1 - R_{n+1}) + \sigma_2(R_2 - R_{n+1}) + \cdots + \sigma_n(R_n - R_{n+1}). \tag{11.24}$$

This can be written

$$K_n = \sum_{k=1}^{n} \beta_{nk}\omega_k, \tag{11.25}$$

where

$$\beta_{nk} = 1 - R_{n+1}/R_k. \tag{11.26}$$

The fact that $R_n \to 0$ monotonely as $n \to \infty$ enables us to show that (11.25) defines a regular method of evaluating series,* that is, $\sum \omega_n = \Lambda$ implies $K_n \to \Lambda$. This completes the proof of Lemma 11.2 and hence the proof of Theorem 11.1.

It is impossible to strengthen Theorem 11.1 by proving that $A \cdot P \supset AP$, even when P is the arithmetic mean transformation C_1. In fact, we prove the following theorem:

THEOREM 11.3. *Corresponding to each transformation of the form*

$$(P) \qquad T_n = (p_1 s_1 + p_2 s_2 + \cdots + p_n s_n)/P_n,$$

where $p_n \neq 0$, $P_n = p_1 + \cdots + p_n > 0$ *for each* $n = 1, 2, \cdots$, *there is a regular transformation*

$$(A) \qquad S_n = \sum_{k=1}^{\infty} a_{nk} s_k,$$

with $a_{nk} \geqq 0$, *such that* $A \cdot P$ *does not include* AP.

Let $\rho_1, \rho_2, \cdots$ denote in order the odd primes. For each $n = 1, 2, 3, \cdots$ let

$$(11.30) \qquad \begin{aligned} a_{nk} &= 0, & k &\neq 2\rho_n, 2\rho_n^2, 2\rho_n^3, \cdots \\ &= 2^{-\sigma}, & k &= 2\rho_n^{\sigma}, & \sigma &= 1, 2, \cdots. \end{aligned}$$

The double series, of which U_n is the sum by rows and V_n the sum by columns, becomes, after removing rows of zeros and factoring the remaining rows

$$(11.31) \qquad \begin{aligned} &(2P_\tau)^{-1}(p_1 s_1 + \cdots + p_{2\tau-1}s_{2\tau-1} + p_{2\tau}s_{2\tau} + 0 + \cdots \\ &+ (4P_{\tau^2})^{-1}(p_1 s_1 + \cdots + p_{2\tau-1}s_{2\tau-1} + p_{2\tau}s_{2\tau} + \cdots + p_{2\tau^2}s_{2\tau^2} + \cdots \\ &+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots , \end{aligned}$$

where $\tau = \rho_n$. For each $n = 1, 2, \cdots$, let

$$(11.32) \qquad s_{2\tau^\sigma} = 1/2^\sigma P_{\tau^\sigma} p_{2\tau^\sigma}, \qquad \sigma = 1, 2, \cdots,$$

and let $s_{2k} = 0$ when k is not of the form p_n^σ. For each k, let

$$(11.33) \qquad s_{2k-1} = -\, p_{2k}s_{2k}/p_{2k-1}.$$

We now have a complete and unique determination of a regular matrix a_{nk} and a sequence s_k. For each n, the series (11.31) converges by rows to 0 and fails to converge by columns. Therefore the sequence s_k is summable AP to 0 and is non-summable $A \cdot P$.

* Carmichael, Bulletin of the American Mathematical Society, vol. 25 (1918), pp. 97–131.

12. **Comparison of AB with $A'B'$ and $A\cdot B$ with $A'\cdot B'$.** It is trivially easy to see that if A, A', B, B' are regular transformations and there is an index n_0 such that $a_{nk}=a'_{nk}$ and $b_{nk}=b'_{nk}$ when $n\geqq n_0$, then $A\sim A'$ and $B\sim B'$. A comparison of the two methods AB and $A'B'$, or the two methods $A\cdot B$ and $A'\cdot B'$ is not so simple. We can however prove the following theorem:

THEOREM 12.1. *Let*

$$(A),\ (A') \qquad S_n = \sum_{k=1}^{\infty} a_{nk}s_k, \qquad S'_n = \sum_{k=1}^{\infty} a'_{nk}s_k,$$

$$(B),\ (B') \qquad T_n = \sum_{k=1}^{\infty} b_{nk}s_k, \qquad T'_n = \sum_{k=1}^{\infty} b'_{nk}s_k,$$

be four regular methods of summability, and let an index N exist such that $a_{nk}=a'_{nk}$, and $b_{nk}=b'_{nk}$, for $n\geqq N$.

Then the two methods of summability

$$(AB) \qquad U_n = \sum_{p=1}^{\infty} a_{np}\sum_{k=1}^{\infty} b_{pk}s_k,$$

$$(A'B') \qquad U'_n = \sum_{p=1}^{\infty} a'_{np}\sum_{k=1}^{\infty} b'_{pk}s_k$$

are consistent, and the two methods

$$(A\cdot B) \qquad V_n = \sum_{k=1}^{\infty}\sum_{p=1}^{\infty} a_{np}b_{pk}s_k$$

$$(A'\cdot B') \qquad V'_n = \sum_{k=1}^{\infty}\sum_{p=1}^{\infty} a'_{np}b'_{pk}s_k$$

are consistent.

We prove first that (AB) and $(A'B')$ are consistent. Suppose s_k is a sequence summable AB to L and summable $A'B'$ to L' so that $U_n\to L$, $U'_n\to L'$. Then

$$T_p = \sum_{k=1}^{\infty} b_{pk}s_k, \qquad T'_p = \sum_{k=1}^{\infty} b'_{pk}s_k$$

must exist for each $p=1, 2, \cdots$. Letting o_α denote generically quantities depending on α which converge to 0 as $\alpha\to\infty$, we find

$$\begin{aligned} U_n &= \sum_{p=1}^{N-1} a_{np}T_p + \sum_{p=N}^{\infty} a_{np}\left[\sum_{k=1}^{N-1} b_{pk}s_k + \sum_{k=N}^{\infty} b_{pk}s_k\right] \\ &= o_n + \sum_{p=N}^{\infty} a_{np}\left[o_p + \sum_{k=N}^{\infty} b_{pk}s_k\right] \\ &= o_n + \sum_{p=N}^{\infty} a_{np}\sum_{k=N}^{\infty} b_{pk}s_k. \end{aligned}$$

Likewise

$$U_n' = o_n + \sum_{p=N}^{\infty} a_{np}' \sum_{k=N}^{\infty} b_{pk}' s_k.$$

But over the ranges of summation in this series, we have, when $n \geqq N$, $a_{np}' = a_{np}$ and $b_{pk}' = b_{pk}$. Hence $U_n = o_n + U_n'$, and it follows that $L = L'$. Thus AB and $A'B'$ are consistent.

To prove that $A \cdot B$ and $A' \cdot B'$ are consistent, let s_k be a sequence summable $A \cdot B$ to Λ and summable $A'B'$ to Λ' so that $V_n \to \Lambda$, $V_n' \to \Lambda'$. Since $a_{np}' = a_{np}$ when $n \geqq N$, we can write for $n > N$

$$V_n = \sum_{k=1}^{\infty} \left(\sum_{p=1}^{N} a_{np} b_{pk} s_k + \sum_{p=N+1}^{\infty} a_{np} b_{pk} s_k \right),$$

$$V_n' = \sum_{k=1}^{\infty} \left(\sum_{p=1}^{N} a_{np} b_{pk}' s_k + \sum_{p=N+1}^{\infty} a_{np} b_{pk}' s_k \right).$$

Since $b_{pk}' = b_{pk}$ when $p > N$, it follows that when $n \geqq N$

$$V_n - V_n' = \sum_{k=1}^{\infty} \sum_{p=1}^{N} a_{np}(b_{pk} - b_{pk}') s_k.$$

An application of the following lemma with $\delta_{pk} = (b_{pk} - b_{pk}')s_k$ shows that $V_n - V_n' \to 0$ as $n \to \infty$ and hence that $\Lambda = \Lambda'$ and that $A \cdot B$ and $A' \cdot B'$ are consistent.

LEMMA 12.2. *If*

$$\lim_{n \to \infty} a_{np} = 0, \qquad p = 1, 2, \cdots, N, \tag{12.21}$$

and

$$\Delta_n = \sum_{k=1}^{\infty} \sum_{p=1}^{N} a_{np} \delta_{pk} \tag{12.22}$$

exists for each sufficiently great n (say $n \geqq n_0$), then

$$\lim_{n \to \infty} \Delta_n = 0. \tag{12.23}$$

We prove this lemma by induction, considering first the case where $N = 1$ and accordingly

$$\Delta_n = \sum_{k=1}^{\infty} a_{n1} \delta_{1k}, \qquad n \geqq n_0.$$

If $a_{n1}=0$ for all sufficiently great n, then obviously $\Delta_n \to 0$. Otherwise we can choose $n \geqq n_0$ such that $a_{n1} \neq 0$ and conclude existence of

$$B_1 = \sum_{k=1}^{\infty} \delta_{1k},$$

so that

$$\Delta_n = a_{n1} B_1.$$

Since $a_{n1} \to 0$, it follows that $\Delta_n \to 0$. Thus the lemma holds when $N=1$.

In case $N=2$, we have

$$\Delta_n = \sum_{k=1}^{\infty} (a_{n1}\delta_{1k} + a_{n2}\delta_{2k}), \qquad n \geqq n_0. \tag{12.24}$$

Suppose two multipliers μ_1 and μ_2, not both zero, exist such that

$$\mu_1 a_{n1} + \mu_2 a_{n2} = 0, \qquad n > n_0.$$

Then supposing $\mu_2 \neq 0$ (the case $\mu_1 \neq 0$ is analogous) and putting $\rho = -\mu_1/\mu_2$ we have

$$a_{n2} = \rho a_{n1}, \qquad n > n_0,$$

and substitution in (12.24) gives

$$\Delta_n = \sum_{k=1}^{\infty} a_{n1}(\delta_{1k} + \rho\delta_{2k}).$$

Now $\Delta_n \to 0$ follows from truth of the lemma for $N=1$.

If multipliers μ_1 and μ_2 do not exist as above, then we can choose n_1 and n_2 such that $n_0 < n_1 < n_2$ and

$$\begin{vmatrix} a_{n_1 1} & a_{n_1 2} \\ a_{n_2 1} & a_{n_2 2} \end{vmatrix} \neq 0. \tag{12.25}$$

From

$$\Delta_{n_1} = \sum_{k=1}^{\infty} (a_{n_1 1}\delta_{1k} + a_{n_1 2}\delta_{2k}),$$

$$\Delta_{n_2} = \sum_{k=1}^{\infty} (a_{n_2 1}\delta_{1k} + a_{n_2 2}\delta_{2k}),$$

we conclude existence of

$$a_{n_2 2}\Delta_{n_1} - a_{n_1 2}\Delta_{n_2} = \sum_{k=1}^{\infty} (a_{n_1 1}a_{n_2 2} - a_{n_1 2}a_{n_2 1})\delta_{1k}.$$

This relation and (12.25) enable us to conclude convergence of the first, and similarly we conclude convergence of the second of the series

$$B_1 = \sum_{k=1}^{\infty} \delta_{1k}, \qquad B_2 = \sum_{k=1}^{\infty} \delta_{2k}.$$

Therefore we can write (12.24) in the form

$$\Delta_n = a_{n1}B_1 + a_{n2}B_2,$$

and $\Delta_n \to 0$ follows from $a_{n1}\to 0$, $a_{n2}\to 0$. Thus the lemma holds when $N=2$.

Assuming that the lemma holds when $N<R$, we can show that

$$\Delta_n = \sum_{k=1}^{\infty} \sum_{p=1}^{R} a_{np}\delta_{pk} \to 0$$

in the case where multipliers $\mu_1, \mu_2, \cdots, \mu_R$ not all zero exist such that

$$\mu_1 a_{n1} + \mu_2 a_{n2} + \cdots + \mu_R a_{nR} = 0, \qquad n > n_0;$$

and also in the alternative case where $n_1 < n_2 < \cdots < n_R$ all exist greater than n_0 and such that the determinant

$$\det \left| a_{n_\alpha \beta} \right|, \qquad \alpha, \beta = 1, \cdots, R,$$

does not vanish. The methods are analogous to those of our proof for the case $N=2$. This completes the proof of Lemma 12.2 and hence of Theorem 12.1.

The reader may naturally be disgruntled by the conclusions of Theorem 12.1; the theorem would be more satisfying if we could replace "consistent" by "equivalent," but we cannot do this. It is clear that under the hypotheses of the theorem AB and $A'B$ are equivalent; and $A\cdot B$ and $A'\cdot B$ are equivalent. But AB and AB' (or $A\cdot B$ and $A\cdot B'$) need not be equivalent.

For an example, put $a_{n1}=1/n$, $n=1, 2, \cdots$; $a_{nn}=1-1/n$, $n=2, 3, \cdots$; and $a_{nk}=0$ when $k\neq 1, n$. Let $b_{1k}=1/2^p$ when k has the form $4p-3$, and $b_{1k}=0$ otherwise. Let $b'_{1k}=1/2^p$ when k has the form $4p-1$ and $b'_{1k}=0$ otherwise. Let for each $n>1$

$$\begin{aligned} b_{nk} = b'_{nk} &= 1, \quad k = n, \\ &= 0, \quad \text{otherwise}. \end{aligned}$$

The sequence $s_k=[(-2)^p$ when $k=4p-1$ and 0 otherwise$]$ is summable AB and $A\cdot B$ to 0 but is non-summable AB' and $A\cdot B'$. The sequence $s_k=[(-2)^p$ when $k=4p-3$ and 0 otherwise$]$ is summable AB' and $A\cdot B'$ to 0 but is non-summable AB and $A\cdot B$. Thus AB and AB' have overlapping convergence fields, as do $A\cdot B$ and $A\cdot B'$.

13. **Multiple products.** In terms of three methods A, B, C of summability, we can define five types of products: ABC, $A(B\cdot C)$ $(A\cdot B)C$, $A\cdot(B\cdot C)$ and $(A\cdot B)\cdot C$. It is easy to show that the last two methods are equivalent. It follows easily from Theorem 7.1 that each other pair selected from the five products may be inconsistent, even though A, B, C are regular and $a_{nk}\geqq 0$, $b_{nk}\geqq 0$, and $c_{nk}\geqq 0$.

14. **Kernel transformations.** Just as matrices (a_{nk}) serve to define generalizations of $\lim_{n\to\infty}s_n$, so also kernels $a(x, t)$ serve to define generalizations of $\lim_{t\to\infty}s(t)$. The $\mathcal{A}$ transform of a function $s(t)$ is, if it exists, given by

$$(\mathcal{A}) \qquad S(x) = \int_0^\infty a(x, t)s(t)dt,$$

and $s(t)$ is summable $\mathcal{A}$ to L if $S(x)\to L$ as $x\to\infty$.

Since the transformation $\mathcal{A}$ becomes essentially a matrix transformation of sequences when we put $s(t)=s_k$ when $k-1\leqq t<k$ and $a(x, t)=a_{nk}$ when $n-1\leqq x<n$, $k-1\leqq t<k$, so that $S(x)=S_n$ when $n-1\leqq x<n$, it follows that some of our results have immediate application to kernel transformations. We mention only the fact that the *iteration product*

$$(\mathcal{A}\mathcal{B}) \qquad U(x) = \int_0^\infty a(x, \alpha)d\alpha \int_0^\infty b(\alpha, t)s(t)dt$$

and the *composition product*

$$(\mathcal{A}\cdot\mathcal{B}) \qquad V(x) = \int_0^\infty \left\{\int_0^\infty a(x, \alpha)b(\alpha, t)d\alpha\right\} s(t)dt$$

may represent inconsistent methods of summability of functions, even though $\mathcal{A}$ and $\mathcal{B}$ are regular and the kernels $a(x, t)$ and $b(x, t)$ are everywhere nonnegative. It thus appears that a formal change of order of integration in a right member above may not only produce a meaningless integral but may actually produce a wrong answer.

CORNELL UNIVERSITY,
ITHACA, N. Y.

THE GEOMETRY OF WHIRL SERIES*

BY

JOHN DE CICCO

INTRODUCTION

In this paper, we shall give some results in addition to those given in a paper, called *The geometry of isogonal and equi-tangential series* by Kasner.†

We begin by considering certain simple operations or transformations on the oriented lineal elements of the plane. A *turn* T_α converts each element into one having the same point and a direction making a fixed angle α with the original direction. By a *slide* S_k the line of the element remains the same and the point moves along the line a fixed distance k. These transformations together generate a continuous group of three parameters which is called the group of *whirl* transformations.‡

Applying a turn T_α to the tangential elements of a union produces an *isogonal series* while a slide S_k produces an *equi-tangential series*. When we apply a whirl to the tangential elements of a union we obtain a *whirl series*.

Kasner has proved that (1) *any element transformation which converts every isogonal series into an isogonal series must be the product of a conformal transformation by a turn. He also has proved that* (2) *any element transformation which carries every equi-tangential series into an equi-tangential series must be the product of an equi-long transformation by a magnification by a slide.*

We shall derive certain generalizations of the above results, two of which are the following: *Any element transformation which carries every union into a whirl series must be the product of a contact transformation by a whirl. Any element transformation which converts every whirl series into a whirl series must be the product of a rigid motion by a magnification by a whirl.*

We also find that, *for an arbitrary element transformation, the maximum number of unions, isogonal, equitangential and whirl series which are transformed into whirl series are respectively* ∞^4, ∞^5, ∞^5, ∞^7.

I. THE DIFFERENTIAL EQUATION OF ALL WHIRL SERIES

For the analytic representation of an element, it will be found convenient to use two systems of coordinates called the cartesian and the hessian co-

* Presented to the Society, April 10, 1936; received by the editors April 16, 1937.

† These Transactions, vol. 42 (1937), pp. 94–106.

‡ Kasner, *The group of turns and slides and the geometry of turbines*, American Journal of Mathematics, vol. 33 (1911), pp. 193–202.

ordinate systems respectively. The cartesian coordinates of an element E are (x, y, θ) where (x, y) are the cartesian coordinates of the point of E and $\theta = \arc\tan p$ is the inclination of the line of E. The hessian coordinates of an element E are (u, v, w) where v is the length of the perpendicular from the origin to the line of E, u is the angle between the perpendicular and the initial line, and w is the distance between the foot of the perpendicular and the point of E.

The equations of the slide S_k are

$$U = u, \qquad V = v, \qquad W = w + k.$$

The equations of the turn T_α are

$$U = u + \alpha, \qquad V = v \cos \alpha + w \sin \alpha, \qquad W = -v \sin \alpha + w \cos \alpha.$$

Since any whirl transformation may be given in the form $W = T_\beta S_k T_\alpha$, it follows that the equations of the whirl $W = T_\beta S_k T_\alpha$ are

$$\text{(1)} \qquad \begin{cases} U = u + \alpha + \beta, \qquad V = v \cos(\alpha+\beta) + w \sin(\alpha+\beta) + k \sin \beta, \\ W = -v \sin(\alpha+\beta) + w \cos(\alpha+\beta) + k \cos\beta. \end{cases}$$

From (1), it is found that any whirl series is given by the equations

$$\text{(2)} \qquad \begin{cases} V = v(U-\alpha-\beta)\cos(\alpha+\beta) + v'(U-\alpha-\beta)\sin(\alpha+\beta) + k\sin\beta, \\ W = -v(U-\alpha-\beta)\sin(\alpha+\beta) + v'(U-\alpha-\beta)\cos(\alpha+\beta) + k\cos\beta, \end{cases}$$

where $(u, v(u), v'(u))$ are the elements of the fundamental union, and α, k, β are the constant parameters of the whirl transformation.

THEOREM 1. *In hessian coordinates the necessary and sufficient condition that the series $V = V(U)$, $W = W(U)$ be a whirl series is that the functions V and W satisfy the equation of third order*

$$\text{(3)} \qquad \frac{d}{dU}\,\frac{V'' - W'}{V' + W''} = 0.$$

In cartesian coordinates the necessary and sufficient condition that the series $X = X(\theta)$, $Y = Y(\theta)$ be a whirl series is that the functions X and Y satisfy the equation

$$\text{(4)} \qquad \frac{d}{d\theta}\left(\frac{-X' + Y''}{X'' + Y'}\right) = 1 + \left(\frac{-X' + Y''}{X'' + Y'}\right)^2.$$

If in hessian coordinates, $V = V(U)$, $W = W(U)$ is a whirl series, then it may be given by the equations (2) which obviously satisfy (3).

Let now the series $V = V(U)$, $W = W(U)$ be such that the functions V

and W satisfy (3). By integrating (3) with respect to U and writing the constant of integration as tan $(\alpha+\beta)$, we obtain the equation

$$(V'' - W')\cos(\alpha+\beta) - (V' + W'')\sin(\alpha+\beta) = 0,$$

which upon a second integration yields

$$V'\cos(\alpha+\beta) - W'\sin(\alpha+\beta) = V\sin(\alpha+\beta) + W\cos(\alpha+\beta) - k\cos\alpha,$$

where k is a constant. From this equation, it is seen that the equations

$$\begin{aligned} u &= U - \alpha - \beta, \\ v &= V\cos(\alpha+\beta) - W\sin(\alpha+\beta) + k\sin\alpha, \\ w &= V\sin(\alpha+\beta) + W\cos(\alpha+\beta) - k\cos\alpha \end{aligned}$$

represent a union. Solving them for U, V, W, and expressing V, W in terms of U, we are led to (2) and thus conclude that our given series is a whirl series. The proof for cartesian coordinates is similar.

COROLLARY. *In hessian coordinates, the necessary and sufficient condition that the series $U=U(t)$, $V=V(t)$, $W=W(t)$ be a whirl series is that the functions U, V, and W satisfy the equation*

$$\frac{d}{dt}\left(\frac{U_tV_{tt} - V_tU_{tt} - W_tU_t^2}{V_tU_t^2 + U_tW_{tt} - W_tU_{tt}}\right) = 0. \tag{5}$$

In cartesian coordinates, the necessary and sufficient condition that the series $X=X(t)$, $Y=Y(t)$, $\theta=\theta(t)$ be a whirl series is that the functions X, Y, and θ satisfy the equation

$$\frac{dM}{dt} = \theta_t(1 + M^2), \tag{6}$$

where

$$M = \frac{-X_t\theta_t^2 + \theta_tY_{tt} - Y_t\theta_{tt}}{\theta_tX_{tt} - X_t\theta_{tt} + Y_t\theta_t^2}.$$

II. UNIONS INTO WHIRL SERIES

THEOREM 2. *Under any element transformation, there exists in general a four-parameter family of unions which are transformed into whirl series. Any element transformation which converts every union into a whirl series must be the product of a contact transformation by a whirl.*

By any element transformation, the elements of a union

$$u, \qquad v = v(u), \qquad w = v'(u),$$

become the elements

$$U = \phi(u, v, v'), \qquad V = \psi(u, v, v'), \qquad W = \chi(u, v, v').$$

If this series is a whirl series, then by (5) we must have

$$\frac{d}{du}\left\{\frac{\dfrac{d\phi}{du}\dfrac{d^2\psi}{du^2} - \dfrac{d\psi}{du}\dfrac{d^2\phi}{du^2} - \dfrac{d\chi}{du}\left(\dfrac{d\phi}{du}\right)^2}{\dfrac{d\psi}{du}\left(\dfrac{d\phi}{du}\right)^2 + \dfrac{d\phi}{du}\dfrac{d^2\chi}{du^2} - \dfrac{d\chi}{du}\dfrac{d^2\phi}{du^2}}\right\} = 0. \tag{7}$$

Upon substituting the values of the derivatives of ϕ, ψ, and χ into (7), we have

$$\frac{d}{du}\left(\frac{A + Bv'' + Cv''^2 + Dv''^3 + Ev'''}{K + Lv'' + Mv''^2 + Nv''^3 + Pv'''}\right) = 0, \tag{8}$$

where $A, B, C, D, E, K, L, M, N, P$ are functions of u, v, v' only.

Since (8) is a differential equation of the fourth order in v, it follows that there is in general a four-parameter family of unions which under the above element transformation are carried into whirl series. All the different cases in the preceding paper have been discussed in a paper by Kasner and De Cicco called *The classification of element transformations by isogonal and equitangential series*, published in the Proceedings of the National Academy of Sciences, vol. 24 (1938), no. 1, pp. 34–38. The different cases of this paper will be considered in a later paper by the author.

Our next problem is to determine all element transformations which convert any union into a whirl series. Then (8) is an identity, and the coefficient of v'''' in (8) is zero, whence

$$\frac{A + Bv'' + Cv''^2 + Dv''^3}{K + Lv'' + Mv''^2 + Nv''^3} = \frac{E}{P}.$$

Since this relation is satisfied for every union, the ratios A/K, B/L, C/M, D/N, E/P have a common value $\lambda(u, v, v')$ and (8) becomes

$$\frac{d\lambda}{du} = \lambda_u + v'\lambda_v + v''\lambda_{v'} = 0.$$

Hence λ is a constant, say $\tan\alpha$, and we obtain

$$\frac{A}{K} = \frac{B}{L} = \frac{C}{M} = \frac{D}{N} = \frac{E}{P} = \tan\alpha. \tag{9}$$

Of the functions $A, B, C, D, E, K, L, M, N, P$, we shall need the explicit expressions for the following ones:

$$
(10)\quad \begin{cases}
A = (\phi_u + v'\phi_v)(\psi_{uu} + 2v'\psi_{uv} + v'^2\psi_{vv}) \\
\qquad - (\psi_u + v'\psi_v)(\phi_{uu} + 2v'\phi_{uv} + v'^2\phi_{vv}) - (\chi_u + v'\chi_v)(\phi_u + v'\phi_v)^2, \\
K = (\psi_u + v'\psi_v)(\phi_u + v'\phi_v)^2 + (\phi_u + v'\phi_v)(\chi_{uu} + 2v'\chi_{uv} + v'^2\chi_{vv}) \\
\qquad - (\chi_u + v'\chi_v)(\phi_{uu} + 2v'\phi_{uv} + v'^2\phi_{vv}), \\
D = \phi_{v'}\psi_{v'v'} - \psi_{v'}\phi_{v'v'} - \chi_{v'}\phi_{v'}^2, \\
N = \psi_{v'}\phi_{v'}^2 + \phi_{v'}\chi_{v'v'} - \chi_{v'}\phi_{v'v'}, \\
E = \psi_{v'}(\phi_u + v'\phi_v) - \phi_{v'}(\psi_u + v'\psi_v), \\
P = \chi_{v'}(\phi_u + v'\phi_v) - \phi_{v'}(\chi_u + v'\chi_v).
\end{cases}
$$

We observe that our element transformation may be expressed as the product of some other transformation

$$U = \Phi(u, v, w), \qquad V = \Psi(u, v, w), \qquad W = \mathrm{X}(u, v, w),$$

by the turn T_α, where α is the constant angle of (9). Then we have

$$\phi = \Phi + \alpha, \qquad \psi = \Psi \cos \alpha + \mathrm{X} \sin \alpha, \qquad \chi = -\Psi \sin \alpha + \mathrm{X} \cos \alpha.$$

It is clear that the equation $E/P = \tan \alpha$ may be written as

$$\frac{\dfrac{\psi_{v'}}{\phi_{v'}} - \dfrac{\psi_u + v'\psi_v}{\phi_u + v'\phi_v}}{\dfrac{\chi_{v'}}{\phi_{v'}} - \dfrac{\chi_u + v'\chi_v}{\phi_u + v'\phi_v}} = \tan \alpha.$$

Substituting the equivalent expressions in terms of Φ, Ψ, X and simplifying, we obtain the relation

$$(11)\qquad \frac{\Psi_u + v'\Psi_v}{\Phi_u + v'\Phi_v} = \frac{\Psi_{v'}}{\Phi_{v'}}$$

not containing α.

It is readily seen that the equation $D/N = \tan \alpha$ may be written as

$$\frac{\dfrac{\partial}{\partial v'}\left(\dfrac{\psi_{v'}}{\phi_{v'}} - \chi\right)}{\dfrac{\partial}{\partial v'}\left(\psi + \dfrac{\chi_{v'}}{\phi_{v'}}\right)} = \tan \alpha.$$

Substituting the equivalent expressions in terms of Φ, Ψ, X, we obtain eventually the equation

$$\frac{\partial}{\partial v'}\left(\frac{\Psi_{v'}}{\Phi_{v'}} - \mathrm{X}\right) = 0,$$

which upon integration with respect to v' becomes

$$\frac{\Psi_{v'}}{\Phi_{v'}} = \mathrm{X} + \mu(u, v), \tag{12}$$

where μ is a function of u and v only.

The equation $A/K = \tan\alpha$ can be put into the form

$$\frac{\left(\dfrac{\partial}{\partial u} + v'\dfrac{\partial}{\partial v}\right)\left(\dfrac{\psi_u + v'\psi_v}{\phi_u + v'\phi_v} - \chi\right)}{\left(\dfrac{\partial}{\partial u} + v'\dfrac{\partial}{\partial v}\right)\left(\psi + \dfrac{\chi_u + v'\chi_v}{\phi_u + v'\phi_v}\right)} = \tan\alpha.$$

Substituting into this equation the equivalent expressions in terms of Φ, Ψ, X we obtain, on simplification

$$\left(\frac{\partial}{\partial u} + v'\frac{\partial}{\partial v}\right)\left[\frac{\Psi_u + v'\Psi_v}{\Phi_u + v'\Phi_v} - \mathrm{X}\right] = 0. \tag{13}$$

From (11), (12), and (13), we have

$$\left(\frac{\partial}{\partial u} + v'\frac{\partial}{\partial v}\right)\mu(u, v) = \mu_u + v'\mu_v = 0.$$

Therefore μ is a constant k.

We therefore conclude that the functions Φ, Ψ, X satisfy the equations

$$\frac{\Psi_u + v'\Psi_v}{\Phi_u + v'\Phi_v} = \frac{\Psi_{v'}}{\Phi_{v'}} = \mathrm{X} + k. \tag{14}$$

Hence it follows that our given transformation is the product of a contact transformation by a slide by a turn, that is, it is the product of a contact transformation by a whirl. That this condition is sufficient is obvious.

III. Equi-tangential series into whirl series

Theorem 3. *Under any element transformation, there exists in general a five-parameter family of equi-tangential series which are transformed into whirl series. Any contact transformation which carries every equi-tangential series into a whirl series must carry every equi-tangential series into an equi-tangential series; that is, it must be the product of an equi-long transformation by a magnification. Any element transformation which converts every equi-tangential series into a whirl series must be the product of an equi-long transformation by a magnification by a whirl.*

By any element transformation the elements of an equi-tangential series u, $v=v(u)$, $w=v'(u)+k$, become the elements

$$U = \phi(u, v, w), \qquad V = \psi(u, v, w), \qquad W = \chi(u, v, w).$$

If this series is a whirl series, then by (5) we have

$$\frac{d}{du}\left[\frac{\dfrac{d\phi}{du}\dfrac{d^2\psi}{du^2} - \dfrac{d\psi}{du}\dfrac{d^2\phi}{du^2} - \dfrac{d\chi}{du}\left(\dfrac{d\phi}{du}\right)^2}{\dfrac{d\psi}{du}\left(\dfrac{d\phi}{du}\right)^2 + \dfrac{d\phi}{du}\dfrac{d^2\chi}{du^2} - \dfrac{d\chi}{du}\dfrac{d^2\phi}{du^2}}\right] = 0. \tag{15}$$

Upon substituting the values of the derivatives of ϕ, ψ, χ into (15), we have

$$\frac{d}{du}\left(\frac{A + Bw' + Cw'^2 + Dw'^3 + Ew''}{K + Lw' + Mw'^2 + Nw'^3 + Pw''}\right) = 0, \tag{16}$$

where $A, B, C, D, E, K, L, M, N, P$ are functions of u, v, w, k only.

Since (16) is a differential equation of the fourth order in v, the complete solution of (16) contains four constants of integration and the constant k of our equi-tangential series. Therefore, there is in general a five-parameter family of equi-tangential series which under the above element transformation are carried into whirl series.

Our next problem is to determine all element transformations which convert any equi-tangential series into a whirl series. Then (16) is an identity, hence the coefficient of w''' is zero. Thus

$$\frac{A + Bw' + Cw'^2 + Dw'^3}{K + Lw' + Mw'^2 + Nw'^3} = \frac{E}{P}.$$

Since this equation is satisfied for every equi-tangential series, the ratios A/K, B/L, C/M, D/N, E/P have a common value $\lambda(u, v, w, k)$, and (16) becomes

$$\frac{d\lambda}{du} = \lambda_u + (w - k)\lambda_v + w'\lambda_w = 0;$$

that is, λ is a function of k only. Thus we obtain the result

$$\frac{A}{K} = \frac{B}{L} = \frac{C}{M} = \frac{D}{N} = \frac{E}{P} = \lambda(k), \tag{17}$$

where λ is a function of k only.

Now since a union is a special case of an equi-tangential series, it follows

by Theorem 2 that the transformation which we are seeking must be the product of a contact transformation by a whirl. It remains therefore to consider the contact transformations which carry any equi-tangential series into a whirl series.

If $\lambda=0$, it follows from (17) that the numerator of the fraction in (16), and hence that of the fraction in (15), vanishes for every equi-tangential series. Thus, the equation

$$\frac{d}{du}\left\{\frac{\dfrac{d\psi}{du}}{\dfrac{d\phi}{du}}-\chi\right\}=0 \tag{18}$$

must be identically satisfied. From (18) it follows that our contact transformation converts every equi-tangential series into an equi-tangential series. Therefore according to Kasner's result our contact transformation must be the product of an equi-long transformation by a magnification.

We shall prove that there is no contact transformation with the desired property for which λ is different from zero. In the first place, our contact transformation can not be an extended line transformation. For any such transformation with the desired property must carry every equi-tangential series into an equi-tangential series, and thus λ must be zero. Therefore for our contact transformation, we must have

$$\chi=\frac{\psi_u+w\psi_v}{\phi_u+w\phi_v}=\frac{\psi_w}{\phi_w}, \tag{19}$$

where $\phi_w\neq 0$ and $\psi_w\neq 0$.

The equation $D/N=\lambda$ has the form

$$\frac{\phi_w\psi_{ww}-\psi_w\phi_{ww}-\chi_w\phi_w^2}{\psi_w\phi_w^2+\phi_w\chi_{ww}-\chi_w\phi_{ww}}=\lambda.$$

Since $\lambda\neq 0$, and the numerator of the fraction vanishes by (19), the denominator must vanish, and we have

$$\frac{\partial}{\partial w}\left(\psi+\frac{\chi_w}{\phi_w}\right)=0. \tag{20}$$

From (19) and (20), we obtain

$$\begin{aligned}\psi&=\alpha+\beta\sin(\phi+\gamma),\\ \chi&=\beta\cos(\phi+\gamma),\end{aligned} \tag{21}$$

where $\alpha,\beta\neq 0,\gamma$ are functions of u and v only.

The equation $E/P=\lambda$ may be written as

$$\frac{\psi_w[\phi_u + (w-k)\phi_v] - \phi_w[\psi_u + (w-k)\psi_v]}{\chi_w[\phi_u + (w-k)\phi_v] - \phi_w[\chi_u + (w-k)\chi_v]} = \lambda$$

and hence, by (19), may be put in the form

$$\frac{1}{\lambda} = \frac{s}{k} + r, \tag{22}$$

where

$$s = \frac{\chi_w(\phi_u + w\phi_v) - \phi_w(\chi_u + w\chi_v)}{\phi_w\psi_v - \psi_w\phi_v}, \qquad r = \frac{\phi_w\chi_v - \chi_w\phi_v}{\phi_w\psi_v - \psi_w\phi_v}. \tag{23}$$

Since the jacobian of the transformation cannot be zero, it follows by (19) that the common denominator of r and s, and also the numerator of s, are each different from zero. Hence, since λ is a function of k only, r and $s\neq 0$ are finite real constants, independent of k.

Substituting the values of ψ and χ into the second of the equations (23), we obtain

$$(\beta_v - r\beta\gamma_v)\cos(\phi+\gamma) - (\beta\gamma_v + r\beta_v)\sin(\phi+\gamma) - r\alpha_v = 0.$$

Since $\phi_w\neq 0$, it is seen that $\beta_v - r\beta\gamma_v = 0$, $\beta\gamma_v + r\beta_v = 0$, $r\alpha_v = 0$. Thus $\beta\neq 0$ and γ are functions of u only.

Substituting the values of ψ and χ into the first of equations (23) and remembering that β and γ are functions of u only, we obtain

$$-\beta_u\cos(\phi+\gamma) + \beta\gamma_u\sin(\phi+\gamma) - s\alpha_v = 0.$$

Since $\phi_w\neq 0$, it is seen that

$$\beta_u = 0, \qquad \beta\gamma_u = 0, \qquad s\alpha_v = 0.$$

Thus β and γ are constants independent of k. Then by (21), χ is a function of ϕ. Hence there is no contact transformation with the desired property for which λ is not zero, and the proof of our theorem is complete.

IV. Isogonal series into whirl series

Theorem 4. *Under any element transformation, there exists in general a five-parameter family of isogonal series which are transformed into whirl series. Any contact transformation which carries every isogonal series into a whirl series must carry every isogonal series into an isogonal series, that is, it must be a conformal transformation. Any element transformation which converts every*

isogonal series into a whirl series must be the product of a conformal transformation by a whirl.

By any element transformation, the elements of an isogonal series

$$x, \qquad y = y(x), \qquad \text{arc tan } p = \text{arc tan } y'(x) + \text{arc tan } k$$

become the elements

$$X = \phi(x, y, p), \qquad Y = \psi(x, y, p), \qquad P = \chi(x, y, p).$$

According to (6), this series is a whirl series if and only if

$$(1 + \chi^2)\frac{df}{dx} = (1 + f^2)\frac{d\chi}{dx}, \tag{24}$$

where

$$f = \frac{-\dfrac{d\phi}{dx}\left(\dfrac{d\theta}{dx}\right)^2 + \dfrac{d\theta}{dx}\dfrac{d^2\psi}{dx^2} - \dfrac{d\psi}{dx}\dfrac{d^2\theta}{dx^2}}{\dfrac{d\theta}{dx}\dfrac{d^2\phi}{dx^2} - \dfrac{d\phi}{dx}\dfrac{d^2\theta}{dx^2} + \dfrac{d\psi}{dx}\left(\dfrac{d\theta}{dx}\right)^2}, \tag{25}$$

and $\theta = \text{arc tan } \chi$.

Upon substituting the values of the derivatives of ϕ, ψ, χ into (25), we have

$$f = \frac{A + Bp' + Cp'^2 + Dp'^3 + Ep''}{K + Lp' + Mp'^2 + Np'^3 + Rp''}, \tag{26}$$

where $A, B, C, D, E, K, L, M, N, R$ are functions of x, y, p, k, only.

Since, by (26), (24) is a differential equation of the fourth order in y, the complete solution of (24) contains four constants of integration and the constant k of our isogonal series. Thus there is in general a five-parameter family of isogonal series which under the above transformation are carried into whirl series.

Our next problem is to determine all element transformations which convert any isogonal series into a whirl series. Then (24) is an identity, hence the coefficient of p''' is zero, whence

$$(K + Lp' + Mp'^2 + Np'^3)E - (A + Bp' + Cp'^2 + Dp'^3)R = 0.$$

Since this equation is also an identity, the ratios $A/K, B/L, C/M, D/N, E/R$ have a common value $\lambda(x, y, p, k)$. Then obviously $f = \lambda$, and from (24) we obtain

$$f = \frac{A}{K} = \frac{B}{L} = \frac{C}{M} = \frac{D}{N} = \frac{E}{R} = \frac{\chi + \alpha}{1 - \alpha\chi}, \tag{27}$$

where α is a function of k only.

Now since a union is a special case of an isogonal series, it follows by Theorem 2 that the transformation which we are seeking must be the product of a contact transformation by a whirl. It remains therefore to consider the contact transformations which carry any isogonal series into a whirl series.

In the first place, let us suppose that our contact transformation is an extended point transformation. Then it must carry every isogonal series into an isogonal series, and therefore according to Kasner's result, it must be a conformal transformation.

Next let us suppose that our contact transformation carries every point into a line. Then it must be of the form

$$\phi = \frac{g_x + pg_y}{f_x + pf_y}, \quad \psi = f\phi - g, \quad \chi = f, \tag{28}$$

where f and g are functions of x and y only. Then (28) must carry every isogonal series into an equi-tangential series. Thus the elements of the union

$$x, \qquad y = y(x), \qquad y' = y'(x)$$

corresponding to the isogonal series

$$x, \qquad y = y(x), \qquad \text{arc tan } p = \text{arc tan } y'(x) + \text{arc tan } k$$

must be carried into the elements of the union

$$X = \phi(x, y, y'), \qquad Y = \psi(x, y, y'), \qquad P = \chi(x, y, y')$$

corresponding to the equi-tangential series

$$X = \phi(x, y, p), \qquad Y = \psi(x, y, p), \qquad P = \chi(x, y, p).$$

If this series is an equi-tangential series, then

$$\frac{d}{dx}\left([\phi(x, y, p) - \phi(x, y, y')]^2 + [\psi(x, y, p) - \psi(x, y, y')]^2\right)^{1/2} = 0.$$

Upon simplifying this equation by means of (28) and setting the coefficient of p' equal to zero, we obtain (since $f_x g_y - f_y g_x \neq 0$)

$$[(1 + pk)f_x + (p - k)f_y]^2 = (k^2 + 1)(f_x + pf_y)^2. \tag{29}$$

Equation (29) is an identity; hence the coefficient of k^2 is zero, whence

$$(pf_x - f_y)^2 = (f_x + pf_y)^2.$$

Since this equation is also an identity, we must have

$$f_x = \pm f_y, \qquad -f_y = \pm f_x,$$

that is, f is a constant. By (28), this is impossible. We have, therefore, proved that there is no contact transformation which carries every point into a line and every isogonal series into a whirl series.

Now we shall prove that there is no contact transformation with the desired property for which $\phi_p \neq 0$, $\psi_p \neq 0$, and $\chi_p \neq 0$. For our contact transformation we must have

$$\chi = \frac{\psi_x + p\psi_y}{\phi_x + p\phi_y} = \frac{\psi_p}{\phi_p}. \tag{30}$$

The equation $E/R = (\chi+\alpha)/(1-\alpha\chi)$ can be written in the form

$$\frac{\psi_p[(1+kp)\chi_x + (p-k)\chi_y] - \chi_p[(1+kp)\psi_x + (p-k)\psi_y]}{\phi_p[(1+kp)\chi_x + (p-k)\chi_y] - \chi_p[(1+kp)\phi_x + (p-k)\phi_y]} = \frac{\chi+\alpha}{1-\alpha\chi},$$

which, when solved for α and combined with (30) gives,

$$\frac{1}{\alpha} = \frac{r}{k} + s, \tag{31}$$

where

$$\begin{aligned} r &= \frac{(1+\chi^2)[\phi_p(\chi_x + p\chi_y) - \chi_p(\phi_x + p\phi_y)]}{\chi_p[-(p\psi_x - \psi_y) + \chi(p\phi_x - \phi_y)]}, \\ s &= \frac{\phi_p(1+\chi^2)(p\chi_x - \chi_y) - \chi_p\{(p\phi_x - \phi_y) + \chi(p\psi_x - \psi_y)\}}{\chi_p[-(p\psi_x - \psi_y) + \chi(p\phi_x - \phi_y)]}. \end{aligned} \tag{32}$$

Since the jacobian of the transformation can not be zero, it follows by (30) that the common denominator of r and s, and also the numerator of r, are each different from zero. Hence, since α is a function of k only, $r \neq 0$ and s are finite real constants and are independent of k. From (31), it is then seen that $\alpha \neq 0$.

The equation $D/N = (\chi+\alpha)/(1-\alpha\chi)$ may be written in the form

$$\frac{-\phi_p\chi_p^2 + 2\chi\psi_p\chi_p^2 + (1+\chi^2)(\chi_p\psi_{pp} - \psi_p\chi_{pp})}{2\chi\phi_p\chi_p^2 + \psi_p\chi_p^2 + (1+\chi^2)(\chi_p\phi_{pp} - \phi_p\chi_{pp})} = \frac{\chi+\alpha}{1-\alpha\chi}.$$

Solving this equation for α, we obtain

$$\alpha = \frac{-(1+2\chi^2)\phi_p + \chi\psi_p + (1+\chi^2)\left[\dfrac{\partial}{\partial p}\left(\dfrac{\psi_p}{\chi_p}\right) - \chi\dfrac{\partial}{\partial p}\left(\dfrac{\phi_p}{\chi_p}\right)\right]}{\chi\phi_p + (1+2\chi^2)\psi_p + (1+\chi^2)\left[\dfrac{\partial}{\partial p}\left(\dfrac{\phi_p}{\chi_p}\right) + \chi\dfrac{\partial}{\partial p}\left(\dfrac{\psi_p}{\chi_p}\right)\right]}. \tag{33}$$

Since $\alpha\neq 0$ and the numerator of the fraction vanishes by virtue of (30), the denominator must vanish, hence we obtain the relation

$$2\chi\phi_p+\frac{\partial}{\partial p}\left(\frac{\phi_p}{\chi_p}\right)+\chi\frac{\partial}{\partial p}\left(\frac{\psi_p}{\chi_p}\right)=0. \tag{34}$$

From (30) and (34), we then obtain

$$\begin{aligned}\phi&=\alpha+\frac{\gamma\chi}{(1+\chi^2)^{1/2}},\\ \psi&=\beta-\frac{\gamma}{(1+\chi^2)^{1/2}},\end{aligned} \tag{35}$$

where α, β, and $\gamma\neq 0$ are functions of x and y only.

In order that (35) be a contact transformation, we must have

$$\beta_x+p\beta_y=\chi(\alpha_x+p\alpha_y)+(\gamma_x+p\gamma_y)(1+\chi^2)^{1/2}. \tag{36}$$

From (32) and (35), we obtain

$$\begin{aligned}r&=\frac{(\alpha_x+p\alpha_y)(1+\chi^2)+(\gamma_x+p\gamma_y)\chi(1+\chi^2)^{1/2}}{(p\beta_x-\beta_y)-\chi(p\alpha_x-\alpha_y)-(p\gamma_x-\gamma_y)(1+\chi^2)^{1/2}},\\ s&=\frac{(p\alpha_x-\alpha_y)+\chi(p\beta_x-\beta_y)}{(p\beta_x-\beta_y)-\chi(p\alpha_x-\alpha_y)-(p\gamma_x-\gamma_y)(1+\chi^2)^{1/2}}.\end{aligned} \tag{37}$$

Since $r\neq 0$, it follows from (36) and (37) that

$$\frac{(\alpha_x+p\alpha_y)+\chi(\beta_x+p\beta_y)}{(p\alpha_x-\alpha_y)+\chi(p\beta_x-\beta_y)}=\frac{s}{r}=t,$$

where t is a constant independent of k. Solving this equation for χ, we obtain

$$\chi=\frac{(\alpha_x+p\alpha_y)-t(p\alpha_x-\alpha_y)}{-(\beta_x+p\beta_y)+t(p\beta_x-\beta_y)}. \tag{38}$$

It is observed that neither the numerator nor the denominator of the fraction in (38) can be zero. For, if either vanished, both must vanish and it would follow that α and β are constants. But, then γ would have to be a constant, by (36), and if α, β, γ were constants, ϕ and ψ, as given by (35), would be functions of χ; and this is impossible.

We shall prove that γ cannot be a constant. For otherwise by (36) and (38) we would have

$$\frac{(\alpha_x+p\alpha_y)-t(p\alpha_x-\alpha_y)}{-(\beta_x+p\beta_y)+t(p\beta_x-\beta_y)}=\frac{\beta_x+p\beta_y}{\alpha_x+p\alpha_y}.$$

Upon setting the term independent of p and the coefficient of p^2 each equal to zero, we obtain

$$\alpha_x(\alpha_x + t\alpha_y) = -\beta_x(\beta_x + t\beta_y),$$
$$\alpha_y(\alpha_y - t\alpha_x) = \beta_y(-\beta_y + t\beta_x).$$

Upon adding these equations, we find that α and β are constants. This proves that γ cannot be a constant.

By (38) we see that χ has an expression of the form

$$\chi = \frac{a + bp}{c + dp}, \tag{39}$$

where a, b, c, and d are functions of x and y only. Also we must have

$$ad - bc \neq 0. \tag{40}$$

For otherwise by (35) and (39), ϕ, ψ, χ would be independent of p.

Since γ is not a constant, it follows by (36) and (39) that

$$(1 + \chi^2)^{1/2} = \frac{[(a + bp)^2 + (c + dp)^2]^{1/2}}{c + dp}$$

is a rational function of p with coefficients which are functions of x and y only. Hence

$$[(a + bp)^2 + (c + dp)^2]^{1/2}$$

must be a perfect square with respect to the letter p. But this can only happen when $ad-bc=0$. By (40) this is impossible. Therefore we have proved that there is no contact transformation with the desired property for which $\phi_p \neq 0$, $\psi_p \neq 0$, and $\chi_p \neq 0$. This completes the proof of our theorem.

V. Whirl series into whirl series

Theorem 5. *Under any element transformation, there exists in general a seven-parameter family of whirl series which are transformed into whirl series. Any contact transformation which carries every whirl series into a whirl series must be the product of a rigid motion by a magnification. Any element transformation which converts every whirl series into a whirl series must be the product of a rigid motion by a magnification by a whirl.*

By any element transformation, the elements of a whirl series

$$u, \qquad v = v(u), \qquad w = w(u),$$

become the elements

$$U = \phi(u, v, w), \qquad V = \psi(u, v, w), \qquad W = \chi(u. v, w).$$

If this series is a whirl series, then by (5), we must have

$$\frac{d}{du}\left[\frac{\dfrac{d\phi}{du}\dfrac{d^2\psi}{du^2}-\dfrac{d\psi}{du}\dfrac{d^2\phi}{du^2}-\dfrac{d\chi}{du}\left(\dfrac{d\phi}{du}\right)^2}{\dfrac{d\psi}{du}\left(\dfrac{d\phi}{du}\right)^2+\dfrac{d\phi}{du}\dfrac{d^2\chi}{du^2}-\dfrac{d\chi}{du}\dfrac{d^2\phi}{du^2}}\right]=0. \tag{41}$$

Now (41) obviously contains w'''. Since w is a combination of the elements $v=v(u)$, $v'=v'(u)$ of the fundamental union, it follows that (41) is a differential equation of the fourth order in v. Hence the complete solution of (41) contains four constants of integration and the three parameters of our whirl series. Thus there is in general a seven-parameter family of whirl series which are carried into whirl series. The remainder of the theorem follows from Theorems 3 and 4.

THEOREM 6. *There is no transformation which carries every isogonal series into an equi-tangential series, or every equi-tangential series into an isogonal series.*

This is an immediate consequence of Theorems 3 and 4.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

AN EXTENSION OF SCHWARZ'S LEMMA*

BY
LARS V. AHLFORS

I. The fundamental inequality

1. To every neighborhood on a Riemann surface there is given a map onto a region of the complex plane. For any two overlapping neighborhoods the corresponding maps are directly conformal.† We agree to denote points on the surface by $\mathfrak{w}$, corresponding values of the local complex parameter by w.

We introduce a Riemannian metric of the form

$$ds = \lambda \mid dw \mid, \tag{1}$$

where the positive function λ is supposed to depend on the particular parameter chosen, in such a way that ds becomes invariant. The metric is regular if λ is of class C_2. In this paper we shall, without mentioning it further, allow λ to become zero, although such points are of course singularities of the metric.

It is well known that the Gaussian curvature of the metric (1) is given by

$$K = -\lambda^{-2} \Delta \log \lambda, \tag{2}$$

and that this expression remains invariant under conformal mappings of the w-plane. We are interested in the case of a metric with negative curvature, bounded away from zero. It is convenient to choose the upper bound of the curvature equal to -4. From (2) it follows that the corresponding λ satisfies the condition

$$\Delta \log \lambda \geqq 4\lambda^2. \tag{3}$$

When we set $u = \log \lambda$ this is equivalent to

$$\Delta u \geqq 4e^{2u}. \tag{4}$$

The hyperbolic metric of the unit circle $|z| < 1$ is defined by

$$d\sigma = (1 - |z|^2)^{-1} |dz| \tag{5}$$

and has the constant curvature -4.

2. Consider now an analytic function $\mathfrak{w} = f(z)$ from the circle $|z| < 1$ to a Riemann surface W. The analyticity is expressed by the fact that every local parameter w is an analytic function of z. To a differential element dz corresponds an element dw whose length does not depend on the direction of dz. The corresponding value of $ds = \lambda |dw| = \lambda_z |dz|$ is therefore uniquely de-

* Presented to the Society, September 8, 1937; received by the editors April 1, 1937.

† For the definition of a Riemann surface see T. Radó, *Über den Begriff der Riemannschen Fläche*, Acta Szeged, vol. 2 (1925).

termined, and we have $\lambda_z = \lambda |w'(z)|$. It is also seen that $u = \log \lambda_z$ satisfies the condition (4) whenever the given metric has a curvature $\leqq -4$. An exception has to be made for the possible zeros of λ_z, corresponding to the zeros of λ and $w'(z)$.

THEOREM A. *If the function $w = f(z)$ is analytic in $|z| < 1$, and if the metric* (1) *of W has a negative curvature $\leqq -4$ at every point, then the inequality*

$$ds \leqq d\sigma \tag{6}$$

will hold throughout the circle.

Proof: Choose an arbitrary $R < 1$ and set $v = \log R(R^2 - |z|^2)^{-1}$ for $|z| < R$. We note that $\Delta v = 4e^{2v}$ and consequently

$$\Delta(u - v) \geqq 4(e^{2u} - e^{2v}). \tag{7}$$

Let us denote by E the open point set in $|z| < R$ for which $u > v$. It is clear that E cannot contain any zeros of λ_z. Hence (7) is valid and shows that $u - v$ is subharmonic in E. It follows that $u - v$ can have no maximum in E and must approach its least upper bound on a sequence tending to the boundary of E. But E can have no boundary points on $|z| = R$, for v becomes positively infinite as z tends to that circle, and at interior boundary points we must have $u - v = 0$, by continuity. A contradiction is thus obtained, unless E is vacuous. The inequality $u \leqq v$ consequently subsists for all points with $|z| < R$, and letting R tend to 1 we find $u \leqq -\log(1 - |z|^2)$ at all points. This is equivalent to (6).

If W is the unit circle and ds its hyperbolic metric, Theorem A is simply the differential form of Schwarz's lemma given by Pick.*

3. Several generalizations of the theorem just proved suggest themselves at once. Since the only thing we need is to prevent the function $u - v$ from having a maximum in E, it is obvious that the assumptions on λ can be considerably weakened, without affecting the validity of the argument. We shall give below two such generalizations which are found to be particularly useful for the applications.

THEOREM A1. *Let λ be continuous and such that at every point, either* (a) *the second derivatives of $u = \log \lambda$ are continuous and satisfy* (4), *or* (b) *it is possible to find two opposite directions n', n'' for which $\partial u/\partial n' + \partial u/\partial n'' > 0$. Then the statement of the previous theorem is still true.*

Opposite directions in the w-plane correspond to opposite directions in the z-plane. At a maximum of $u - v$ we have $\partial u/\partial n \leqq \partial v/\partial n$ in any direction, when-

* An account of all questions related to Schwarz's lemma will be found in R. Nevanlinna, *Eindeutige analytische Funktionen*, Springer, 1936, pp. 45–58.

ever the directional derivative exists. For opposite directions $\partial v/\partial n' + \partial v/\partial n'' = 0$; hence $\partial u/\partial n' + \partial u/\partial n'' \leqq 0$ in case of a maximum. It follows that no maximum can be attained in points satisfying condition (b).

We shall call $ds' = \lambda' |dw|$ a supporting metric of $ds = \lambda |dw|$ at the point w_0 if: (1) $\lambda' = \lambda$ at w_0, (2) λ' is defined and $\leqq \lambda$ in a neighborhood of w_0.

THEOREM A2. *Suppose that λ is continuous, and that it is possible to find a supporting metric, satisfying* (4), *at every point of W. Then the inequality* (6) *still holds.*

If $u - v > 0$ at z_0, then $u' - v$ will also be positive, and consequently subharmonic, in a neighborhood of z_0.* A maximum of $u - v$ will *a fortiori* be a maximum of $u' - v$. Hence $u - v$ can have no maximum in E.

II. SCHOTTKY'S THEOREM

4. As a first application we prove Schottky's theorem with definite numerical bounds.

THEOREM B. *If $f(z)$ is analytic and different from* 0 *and* 1 *in* $|z| < 1$, *then*

$$\log |f(z)| < \frac{1+\theta}{1-\theta}(7 + \overset{+}{\log} |f(0)|) \tag{8}$$

for $|z| \leqq \theta < 1$.†

Let $\zeta_1 = \zeta_1(w)$ map the region outside of the segment $(0, 1)$ onto the exterior of the unit circle, so that $w = \infty$ corresponds to $\zeta_1 = \infty$, $w = 1$ to $\zeta_1 = 1$, and $w = 0$ to $\zeta_1 = -1$. We also set $\zeta_2(w) = \zeta_1(w^{-1})$ and $\zeta_3(w) = \zeta_2(1 - w)$. Clearly these functions define similar maps of the regions outside of the segments $(1, \infty)$ and $(-\infty, 0)$. Explicitly, $\zeta_1(w)$ is obtained from the equation

$$\zeta_1 + \zeta_1^{-1} = 4w - 2. \tag{9}$$

We introduce the coordinates $\rho_1 = |w|$, $\rho_2 = |w - 1|$ and divide the plane into regions Ω_1: $\rho_1 \geqq 1$, $\rho_2 \geqq 1$; Ω_2: $\rho_1 \leqq 1$, $\rho_1 \leqq \rho_2$; Ω_3: $\rho_2 \leqq 1$, $\rho_2 \leqq \rho_1$. The metric

$$ds_i = \frac{|d \log \zeta_i|}{2(4 + \log |\zeta_i|)} = \lambda_i |dw| \tag{10}$$

---

* u' corresponds to λ' as u to λ.

† Schottky's original theorem was purely qualitative. Numerical relations have been studied at great length, notably by Ostrowski (*Studien über den Schottky'schen Satz*, Basel, 1931, and *Asymptotische Abschätzung des absoluten Betrags einer Funktion, die die Werte* 0 *und* 1 *nicht annimmt*, Commentarii Mathematici Helvetici, vol. 5 (1933)), but no simple inequality comparable with (8) has ever been proved.

Added in proof: Numerical bounds of the same order of magnitude are found by A. Pfluger, *Über numerische Schranken im Schottky'schen Satz*, Commentarii Mathematici Helvetici, vol. 7 (1935). His proof depends on the use of modular functions, while ours is strictly elementary.

is readily recognized as the hyperbolic metric of a half-plane with the constant curvature -4. Computing the derivatives $\zeta_i'(w)$ we find

$$\lambda_1^{-1} = 2(\rho_1\rho_2)^{1/2}(4 + \log |\zeta_1|),$$
$$(11) \qquad \lambda_2^{-1} = 2\rho_1\rho_2^{1/2}(4 + \log |\zeta_2|),$$
$$\lambda_3^{-1} = 2\rho_2\rho_1^{1/2}(4 + \log |\zeta_3|).$$

We now set $ds = \lambda |dw|$ with $\lambda = \lambda_i$ in Ω_i. This metric is regular and satisfies condition (3) except at the singular points $0, 1, \infty$ and on the lines separating the regions Ω_i. On these lines λ is still continuous, as seen from (11) and the relations between ζ_1, ζ_2, and ζ_3.

Next we wish to show that condition (b) in Theorem A1 holds on the singular lines. We consider the arc $\rho_1 = 1$, $\rho_2 > 1$ and choose n', n'' as the outer and inner normals of the circle. The required condition is

$$\frac{\partial}{\partial n'} \log \lambda_1 + \frac{\partial}{\partial n''} \log \lambda_2 = \frac{\partial}{\partial n'} \log \frac{\lambda_1}{\lambda_2} > 0.$$

From (11) we obtain

$$\frac{\partial}{\partial n'} \log \frac{\lambda_1}{\lambda_2} = \frac{1}{2} - \frac{\dfrac{\partial}{\partial n'} \log \left| \dfrac{\zeta_1}{\zeta_2} \right|}{4 + \log |\zeta_1|},$$

which is also equal to

$$\frac{1}{2} - 2(4 + \log |\zeta_1|)^{-1} \frac{\partial \Phi_1}{\partial \phi},$$

where $\Phi_1 = \arg \zeta_1$, $\phi = \arg w$. For Φ_1 we have the simple relation $\cos \Phi_1 = \rho_1 - \rho_2$, which for $\rho_1 = 1$ becomes $\cos \Phi_1 = 1 - 2 \sin \phi/2$. Differentiating we find

$$\frac{\partial \Phi_1}{\partial \phi} = \frac{1}{2}\left(1 + \csc \frac{\phi}{2}\right)^{1/2},$$

and by use of the inequalities $\pi/3 \leqq \phi \leqq 5\pi/3$, $|\zeta_1| > 1$, we are finally led to the desired result,

$$\frac{\partial}{\partial n'} \log \frac{\lambda_1}{\lambda_2} > \frac{1}{2} - \frac{3^{1/2}}{4} > 0.$$

By symmetry, the same must be true for the arc $\rho_2 = 1$, $\rho_1 > 1$. The transformation $w' = (1 - w)^{-1}$ takes Ω_1 into Ω_2 and Ω_2 into Ω_3. Since the function λ is invariant under the transformation we conclude at once that condition (b) will hold also on the line separating Ω_2 and Ω_3.

From Theorem A1 we can now conclude that $w = f(z)$ satisfies the differ-

ential inequality $\lambda|dw| \leqq (1-|z|^2)^{-1}|dz|$. Integrating, we find that the shortest distance between the points $f(0)$ and $f(z)$, $|z| = \theta$, measured in the metric $ds = \lambda|dw|$, cannot exceed $[\log (1+\theta)/(1-\theta)]/2$.

The shortest path between the circles $\rho_1 = m$ and $\rho_1 = M$, where $M > m \geqq 2$, is a segment of the negative real axis, whose length is found to be

$$\frac{1}{2} \log \frac{4 + \log |\zeta_1(-M)|}{4 + \log |\zeta_1(-m)|}.$$

To simplify we introduce the lower and upper bounds $|\zeta_1(-M)| \geqq 4M$, $|\zeta_1(-m)| \leqq 5m$. Setting $M = |f(z)|$ and m equal to the greater of the numbers $|f(0)|$ and 2 we obtain

$$4 + \log 4M \leqq \frac{1+\theta}{1-\theta}(4 + \log 5m).$$

Here $\log 5m \leqq \log 10 + \overset{+}{\log} |f(0)| < 3 + \overset{+}{\log} |f(0)|$ and we find

$$4 + \log 4M < \frac{1+\theta}{1-\theta}(7 + \overset{+}{\log} |f(0)|)$$

which is stronger than (8).

III. Bloch's theorem

5. Let $w = f(z)$ be analytic in $|z| < 1$ with $|f'(0)| = 1$. Let $B' = B'(f)$ be the l.u.b. of the radii of all simple (*schlicht*) circles contained in the Riemann surface W generated by $f(z)$. Bloch's theorem is $B = \min B' > 0$. Landau has proved $B > .396$.* Grunsky and Ahlfors proved in a recent paper $B < .472$.†

We show that the method developed in this paper gives an immediate proof of Bloch's theorem with a better lower bound for B. For an arbitrary point $\mathfrak{w}$ on W let $\rho(\mathfrak{w})$ denote the radius of the largest simple circle of center $\mathfrak{w}$ contained in W. It is clear that $\rho(\mathfrak{w})$ is continuous, and equal to zero only at the branch-points. We introduce the metric $ds = \lambda|dw|$ with

$$\lambda = \frac{A}{2\rho^{1/2}(A^2 - \rho)} \qquad (\rho = \rho(\mathfrak{w})) \tag{12}$$

and w denoting the variable of the function plane (not the uniformizing variable). A is a constant satisfying the preliminary condition $A^2 > B'$.

In the neighborhood of a branch-point $\mathfrak{a}$ we have $\rho = |w - a|$. Let n be the multiplicity of $\mathfrak{a}$; then $w_1 = (w-a)^{1/n}$ is a uniformizing variable, and

* E. Landau, *Über den Blochschen Satz und zwei verwandte Weltkonstanten*, Mathematische Zeitschrift, vol. 30 (1929).

† L. V. Ahlfors and H. Grunsky, *Über die Blochsche Konstante*, Mathematische Zeitschrift, vol. 42 (1937). The result was found independently by R. M. Robertson.

the corresponding λ_1 is determined from $\lambda_1|dw_1| = \lambda|dw|$. We obtain $\lambda_1 = n\rho^{1/2-1/n}/2(A^2-\rho)$, and it is seen at once that the metric is regular in case $n=2$ and that λ_1 becomes zero in case $n>2$.

We wish to apply Theorem A2 and therefore look for a supporting metric satisfying the requirements of that theorem. For a regular point $\mathfrak{w}_0$ the surrounding circle of radius $\rho(\mathfrak{w}_0)$ must pass through at least one singularity $\mathfrak{b}$ which is either a branch-point or a boundary point for the surface. We set $\rho' = |w-b|$ and define $\lambda' = A/[2\rho'^{1/2}(A^2-\rho')]$. This metric has the curvature -4 for it is obtained from the hyperbolic metric of a circle by means of the transformation $w' = w^{1/2}$. In all points of our circle we have $\rho \leqq \rho'$ by the definition of ρ. The inequality $\lambda' \leqq \lambda$ is therefore satisfied in a neighborhood of $\mathfrak{w}_0$ if the function $t^{1/2}(A^2-t)$ increases for $t \leqq \rho(\mathfrak{w}_0)$. Under this condition λ' will be a supporting function of λ, for at the center w_0 we have $\lambda' = \lambda$. The function $t^{1/2}(A^2-t)$ is increasing as long as $t<A^2/3$. Consequently all the conditions in Theorem A2 are fulfilled if we suppose that $A^2>3B'$.

Apply the theorem with $z=0$. Using the condition $|dw/dz|_{z=0}=1$ we get

$$A \leqq 2\rho_0^{1/2}(A^2 - \rho_0), \tag{13}$$

where ρ_0 is the radius of the largest simple circle with center at the image of $z=0$. The function in the right member of (13) is increasing, and we can replace ρ_0 by B' obtaining $A \leqq 2B'^{1/2}(A^2-B')$. Letting A tend to $(3B')^{1/2}$ we finally get $B' \geqq 3^{1/2}/4$. This implies that Bloch's constant $B \geqq 3^{1/2}/4 > .433$.

On the other side, if we insert $A^2 = (3B')^{1/2}$ in (13), lower and upper bounds for ρ_0 in terms of B' can be found.

6. Landau has considered a closely related constant L. Let $L' = L'(f)$ be the l.u.b. of the radii of all circles in the w-plane contained in the projection of W, that is, whose values are taken by the function $w=f(z)$, $|f'(0)|=1$. L is defined as the minimum of all such L'. Clearly, $L \geqq B$.

The method employed above is immediately applicable if we choose $\lambda = (2\rho \log C/\rho)^{-1}$. This metric is regular at all branch-points, and when we replace ρ by the distance ρ' from a fixed boundary point, the curvature becomes -4. In order that the function λ' thus obtained be a supporting function it is sufficient that $t \log C/t$ is increasing. This is true for $t<Ce^{-1}$. We therefore choose $C>eL'$, obtaining the inequality $1 \leqq 2L' \log C/L'$ as above. Letting C tend to eL' we find $L' \geqq 1/2$ and hence $L \geqq 1/2$.

This lower bound is the best known. It shows in particular that $L>B$.*

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

* In the other direction R. M. Robinson has proved $L<.544$. This result has not been published.

NORMALITY AND ABNORMALITY IN THE CALCULUS OF VARIATIONS*

BY

G. A. BLISS

Within the past few years a number of papers concerning the problem of Bolza in the calculus of variations have been published which make it possible to carry through the theory of this problem with much simplified assumptions concerning what is called the normality of the minimizing arc. I refer especially to papers by Graves [8],† Hestenes [11, 14, 16], Reid [15], and Morse [13]. These papers and others are also important because they bring the theory of problems of the calculus of variations with variable end points to a stage comparable with that already attained for the more special case in which the end points are fixed.

In the theories of Bolza [1, chap. 11, 12] and Bliss [2] for the problem of Lagrange with fixed end points it was assumed that the minimizing arc considered, extended slightly at both ends, was normal on every sub-interval. Morse [4] showed that the theory could be carried through on the assumption that the arc itself, without extensions, was normal on every sub-interval. The most important case, however, turns out to be the one for which the arc as a whole is normal relative to the problem considered, but not necessarily normal on sub-intervals. Graves proved the necessary condition of Weierstrass for such a normal minimizing arc, and Hestenes deduced further necessary conditions and gave sufficiency proofs for a minimum. The importance of these results is emphasized by the fact that for the very general problem of Mayer, which may be regarded as a sub-case of the problem of Bolza, every minimizing arc is abnormal on every sub-interval, even though the arc as a whole is normal relative to the problem. Thus the problem of Mayer needs a separate treatment, such as was given by Bliss and Hestenes [9, 10], unless one has at his command results equivalent to the recent extensions of the theory of the problem of Bolza mentioned above.

In this paper I am attempting to analyze, more explicitly than has been done before, the meaning of normality and abnormality for the calculus of variations. To do this I have emphasized in §1 below the meaning of normality for the problem of a relative minimum of a function of a finite number of variables. In §2 analogous notions are discussed for problems of the cal-

* Presented to the Society, April 20, 1935; received by the editors April 1, 1937.

† The numbers in brackets here and elsewhere refer to the bibliography at the end of this paper.

culus of variations. From this discussion it will be clear that a normal arc for the problem of Bolza is a non-singular arc of the class in which a minimizing one is sought. The singular arcs of the class are the abnormal ones. They have an enormous variety of types. It is not likely that a general theory can be formulated which would apply to all of them, though one might characterize and study successfully some very general cases.

In the papers of Graves and Hestenes mentioned above there is no explicit assumption concerning normality. The arc studied is assumed only to have a set of multipliers like those which it would have if it were normal for the problem of Bolza considered. In the following pages it will be seen that, though such an arc may be abnormal for the problem originally considered, it is nevertheless normal for a second problem of Bolza obtained from the first by suitably extending the class of arcs in which a minimizing one is sought. Furthermore the properties characterizing a minimizing arc for the original problem are effective for the second, so that the sufficiency theorems of Hestenes for arcs which are normal have as easy consequences those for the abnormal arcs permitted by his hypotheses. This makes possible a number of simplifications in the details of the proofs. It is not to be expected, of course, that new necessary conditions on a minimizing arc can be secured by extending the class of arcs in which a minimizing one is sought. The paper of Graves, therefore, seems to contain results not attainable by considering only normal arcs.

In the introduction to his paper [13] Morse makes a statement concerning priority for the proofs of sufficiency theorems without assumptions of normality which might easily be misunderstood and about which I should like to make the following comments. Hestenes had previously proved, in his paper [11], three sufficiency theorems (Theorems 9:1, 9:3, 9:5) without explicit assumptions of normality, and also a fourth theorem (Theorem 9:4) with normality assumptions still undesirably strong, but weaker than those which had before been used. Reid [15] and Morse [13] showed independently that by means of a further lemma, but aided still essentially by the results of Hestenes, this fourth theorem can be brought to a par with the others. The condition VI′ [11, p. 811] of Theorem 9:4 is analogous to one which I used in the paper [5], and which was originally due to A. Mayer. Its statement involves the notion of conjugate points and is therefore more closely related to the classical conditions of Jacobi for simpler problems than the corresponding conditions of the other theorems. I think it should be understood that the priority comment of Morse is applicable to Theorem 9:4 of Hestenes, but not to the other three theorems of his paper, which are equally important. I may add that the theorems of Hestenes were proved with great originality

and ingenuity while he was my research assistant at the University of Chicago in 1933 [16, p. 543]. When he went away he left a manuscript with me in which the theorems were, at my suggestion, deduced only for normal arcs, the ones which then, as well as now, seemed to me the most important, even though the justification of the arguments of the present paper was at that time missing. This manuscript has since appeared in much modified form in my mimeographed lectures on the problem of Bolza [12]. In his paper [11] Hestenes showed that his methods are also effective for the problem of Bolza in the form adopted by Morse.

1. **Abnormality for minima of functions of a finite number of variables.** The significance of the notion of abnormality in the calculus of variations can be indicated by a study of the theory of the simpler problem of finding, in the set of points $y=(y_1, \cdots, y_n)$ satisfying a system of equations of the form

$$\phi_\beta(y)=0 \ (\beta=1, \cdots, m<n),$$

one which minimizes a function $f(y)$. For a point $y^0=(y_1^0, \cdots, y_n^0)$ near which the functions f and ϕ_β have continuous partial derivatives of at least the second order, and which satisfies the equations $\phi_\beta=0$, we have the following theorems, some of which are, of course, well known.

THEOREM 1:1. *A first necessary condition for $f(y^0)$ to be a minimum is that there exist constants l_0, l_β not all zero such that the derivatives F_{y_i} of the function*

$$F=l_0f+l_\beta\phi_\beta$$

all vanish at y^0.

To prove this we have only to note that the determinants of the matrix

$$\left\| \begin{matrix} f_{y_i}(y^0) \\ \phi_{\beta y_i}(y^0) \end{matrix} \right\|$$

must all vanish. Otherwise, according to well known implicit function theorems, the equations $f(y)=f(y^0)+u$, $\phi_\beta(y)=0$ would have solutions y for negative values of u, and $f(y^0)$ could not be a minimum.

A point y^0 has by definition *order of abnormality equal to q* if there exist q linearly independent sets of multipliers of the form $l_0=0$, l_β having the property of the theorem. When $q=0$ the point y^0 is said to be *normal*. A necessary and sufficient condition for abnormality of order q is evidently that the matrix $\|\phi_{\beta y_i}(y^0)\|$ have rank $m-q$. At a normal point y^0 the multipliers l_0, l_β of the theorem can be divided by l_0 and put into the form $l_0=1$, l_β. In this form they are unique, since the non-vanishing difference of two such sets would be a set of multipliers implying abnormality.

LEMMA 1:1. *If a point y^0 is normal, then for every set of constants η_i $(i=1, \cdots, n)$ satisfying the equations*

$$\phi_{\beta y_i}(y^0)\eta_i = 0 \tag{1:1}$$

there exists a set of functions $y_i(b)$ having continuous second derivatives near $b=0$, satisfying the equations $\phi_\beta=0$, and such that

$$y_i(0)=y_i^0, \quad y_i'(0)=\eta_i.$$

The proof can be made by considering the equations

$$\phi_\beta(y) = 0, \; \phi_\gamma(y) = \phi_\gamma(y^0) + b\zeta_\gamma \quad (\beta = 1, \cdots, m; \gamma = m+1, \cdots, n) \tag{1:2}$$

in which the auxiliary functions $\phi_\gamma(y)$ are selected so that they have continuous second derivatives near y^0 and make the functional determinant $|\phi_{iy_k}(y^0)|$ different from zero, and in which the constants ζ_γ are defined by the equations

$$\phi_{\gamma y_i}(y^0)\eta_i = \zeta_\gamma. \tag{1:3}$$

Equations (1:2) then have solutions $y_i(b)$ with continuous derivatives of at least the second order near $b=0$, and such that $y_i(0)=y_i^0$. By differentiating with respect to b the equations (1:2) with these solutions substituted, we find the equations

$$\phi_{\beta y_i}(y^0)y_i'(0) = 0,$$

$$\phi_{\gamma y_i}(y^0)y_i'(0) = \zeta_\gamma.$$

With equations (1:1) and (1:3) these show that $y_i'(0)=\eta_i$.

THEOREM 1:2. *If y^0 is a normal point and $f(y^0)$ a minimum then the condition*

$$F_{y_iy_k}(y^0)\eta_i\eta_k \geqq 0$$

must hold for every set η_i satisfying the equations (1:1), *where $F=f+l_\beta\phi_\beta$ is the function formed with the unique set of multipliers $l_0=1$, l_β belonging to y^0.*

The conclusion of the theorem is due to the fact that the function $g(b)=f[y(b)]$, formed with the functions $y_i(b)$ of the lemma, must have a minimum at $b=0$. Since

$$\phi_{\beta y_i}[y(b)]y_i'(b) = 0$$

the derivatives of $g(b)$ are seen to have the values

$$g'(b) = f_{y_i}[y(b)]y_i'(b) = F_{y_i}[y(b)]y'(b),$$
$$g''(0) = F_{y_iy_k}(y^0)\eta_i\eta_k,$$

and for $g(0)$ to be a minimum we must have $g''(0) \geqq 0$.

THEOREM 1:3. *If a point y^0 has a set of multipliers $l_0=1$, l_β for which the function $F=f+l_\beta\phi_\beta$ satisfies the conditions*

$$F_{y_i}(y^0) = 0, \qquad F_{y_iy_k}(y^0)\eta_i\eta_k > 0 \tag{1:4}$$

for all sets η_i satisfying the equations

$$\phi_{\beta y_i}(y^0)\eta_i = 0, \tag{1:5}$$

then $f(y^0)$ is a minimum.

This can be proved with the help of Taylor's formula with integral form of remainder. For every point y near y^0 satisfying the equations $\phi_\beta=0$ we have the equations

$$\begin{aligned} f(y) - f(y^0) &= f_{y_i}(y^0)\eta_i + \int_0^1 (1-\theta)f_{y_iy_k}(y')\eta_i\eta_k d\theta, \\ 0 &= \phi_{\beta y_i}(y^0)\eta_i + \int_0^1 (1-\theta)\phi_{\beta y_iy_k}(y')\eta_i\eta_k d\theta, \\ 0 &= \int_0^1 \phi_{\beta y_i}(y_i')\eta_i d\theta, \end{aligned} \tag{1:6}$$

where $y_i' = y_i^0 + \theta(y_i - y_i^0)$, $\eta_i = y_i - y_i^0$. From these we find readily

$$f(y) - f(y^0) = \int_0^1 (1-\theta)F_{y_iy_k}(y')\eta_i\eta_k d\theta.$$

Since the quadratic form in the integrand of the last integral, thought of as a function of independent variables y' and η, is positive for $y'=y^0$ and all sets η satisfying the equations (1:5), it will remain positive for $y'=y^0+\theta(y-y^0)$ and all sets η, including $\eta=y-y^0$, satisfying equations (1:6), provided that y lies in a sufficiently small neighborhood N of the point y^0. Thus we see that for all points y in N satisfying the equations $\phi_\beta=0$ the difference $f(y)-f(y^0)$ is positive.

The last theorem is analogous to the sufficiency theorems of Hestenes in the calculus of variations. In it there is no explicit assumption concerning the normality or abnormality of the point y^0. If y^0 has abnormality of order q, however, let ν be a variable which ranges over a subset of $m-q$ of the numbers $1, \cdots, m$ such that the matrix $\|\phi_{\nu y_i}(y^0)\|$ has rank $m-q$, and let ρ range over the complementary subset. Then we have the following theorem:

THEOREM 1:4. *Let y^0 be a point which satisfies the hypotheses of Theorem* 1:3 *with a set of multipliers $l_0=1$, l_β, and let ν and ρ be variables having the ranges described in the last paragraph. Then y^0 is normal for the modified problem of minimizing the function $g=f+l_\rho\phi_\rho$ in the class of points y satisfying the restricted system of equations $\phi_\nu=0$, and y^0 satisfies the hypotheses of Theorem* 1:3 *for the modified problem with the multipliers $l_0=1$, l_ν. Furthermore if $g(y^0)$ is a minimum for the modified problem, then $f(y^0)$ is a minimum for the original one.*

We see that the point y^0 is normal for the modified problem, since the matrix $\|\phi_{\nu y_i}(y^0)\|$ has rank $m-q$. For the function $F=g+l_\nu\phi_\nu=f+l_\beta\phi_\beta$ of the modified problem the conditions (1:4) are satisfied for all sets η satisfying the equations

$$\phi_{\nu y_i}(y^0)\eta_i = 0, \tag{1:7}$$

since equations (1:5) are linear and have a matrix of coefficients of rank $m-q$ and hence are consequences of equations (1:7). The set of points y satisfying the equations $\phi_\nu=0$ includes the points satisfying the complete system $\phi_\beta=0$ as a subclass in which $g=f$. Hence if $g(y^0)$ is a minimum for the modified problem, the value $f(y^0)=g(y^0)$ must have the same property for the original problem.

From the last theorem it is evident that generality is not lost by proving Theorem 1:3 only for points y^0 which are normal. Such points are, in fact, the non-singular points of the class which satisfy the equations $\phi_\beta=0$. Near each of them there are infinitely many points of the class, as is shown by Lemma 1.1, and the minimum problem near one of them is therefore never trivial. Abnormal points, on the other hand, are the singular points of the class, and may occur in a wide variety of types. For some of these points the minimum problem is trivial, as, for example, in the case of a point y^0, for which $\phi_1=0$, which minimizes the function ϕ_1 in the class of points y satisfying the equations $\phi_2=\cdots=\phi_m=0$. Near such a point y^0 there is no other point satisfying all of the equations $\phi_\beta=0$.

An idea of the great variety of types of abnormal points may be gained by considering the problem of minimizing a function $f(y_1, y_2)$ of two variables in the class of points (y_1, y_2) satisfying a single equation $\phi(y_1, y_2)=0$. The variety of abnormal points possible in this case is at least as great as the variety of singular points of an algebraic curve. The particular example $f=2y_1^2-y_2^2$, $\phi=y_1^2y_2-y_2^3=0$, with minimizing point (0, 0), shows that the condition involving the quadratic form in Theorem 1:3 is not in general necessary for a minimum.

2. **Abnormality in the calculus of variations.** The problem to be considered in this section [12, p. 4] is that of finding in a class of arcs

$$y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2) \tag{2:1}$$

satisfying conditions of the form

$$\phi_\beta(x, y, y') = 0 \qquad (\beta = 1, \cdots, m < n),$$

$$\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \qquad (\mu = 1, \cdots, p \leqq 2n + 2)$$

one which minimizes a sum

$$J = g[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} f(x, y, y')dx.$$

A set of values (x, y, y') and end values $[x_s, y_{is}] = [x_s, y_i(x_s)]$ $(s = 1, 2)$ is said to be *admissible* if it lies interior to a region of such values in which the functions $f, g, \phi_\beta, \psi_\mu$ have continuous derivatives of at least the fourth order, and in which the matrix $\|\phi_{\beta y_i'}\|$ and the matrix of first derivatives of the functions ψ_μ have ranks m and p, respectively. An *admissible arc* C defined by functions of the form (2:1) is one which is continuous and consists of a finite number of sub-arcs with continuously turning tangents, and whose elements (x, y, y') and end values are admissible. When convenient we may represent by $J(C)$, $g(C), \psi_\mu(C)$ the values of these functions determined by the arc C.

The conditions involved in the sufficiency theorems for this problem are the following, the numbering being that which I have often used [see, e.g., 12, chap. 3]:

I. The multiplier rule. A set of multipliers $l_0, l_\beta(x), e_\mu$ for an admissible arc E is a set for which the l_0, e_μ are constants and the functions $l_\beta(x)$, defined on the interval x_1x_2 belonging to E, are continuous except possibly at values of x defining corners of E at which they nevertheless have well-defined forward and backward limits. The arc E satisfies the multiplier rule if there exist constants c_i and multipliers $l_0, l_\beta(x), e_\mu$ such that for $F = l_0 f + l_\beta(x)\phi_\beta$ the equations

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i}dx + c_i, \qquad \phi_\beta = 0$$

are satisfied along E, and furthermore such that the end values of E satisfy the equations

$$[(F - y_i'F_{y_i'})dx + F_{y_i'}dy_i]_1^2 + l_0dg + e_\mu d\psi_\mu = 0, \qquad \psi_\mu = 0 \tag{2:2}$$

identically in the differentials dx_s, dy_{is}.

It has been proved [12, p. 27] that the identically vanishing set of multipliers is the only set having constants l_0, e_μ all zero, or having functions l_0, $l_\beta(x)$ which vanish simultaneously at some value x on the interval x_1x_2.

II_N'. An admissible arc E satisfies the strengthened condition of Weierstrass if for every set of the type (x, y, y', l) in a neighborhood N of those belonging to E the inequality

$$E(x, y, y', l, Y') > 0$$

is satisfied for all admissible sets $(x, y, Y') \neq (x, y, y')$, where

$$E = F(x, y, Y', l) - F(x, y, y', l) - (Y_i' - y_i')F_{y_i'}(x, y, y', l).$$

III$'$. An admissible arc E satisfies the strengthened condition of Clebsch if at every element (x, y, y', l) belonging to E, the inequality

$$F_{y_i'y_k'}(x, y, y', l)\pi_i\pi_k > 0$$

is satisfied for all non-vanishing sets π_i satisfying the equations

$$\phi_{\beta y_i'}(x, y, y')\pi_i = 0.$$

If we represent by q, q_μ the quadratic forms in dx_s, dy_{is} whose coefficients are the second derivatives of the functions g, ψ_μ, respectively, the second variation of J for an extremal arc E with multipliers $l_0=1$, $l_\beta(x)$, e_μ has the value

$$J_2(\xi, \eta) = 2\gamma[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] + \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx$$

in which

$$2\omega = F_{y_iy_k}\eta_i\eta_k + 2F_{y_iy_k'}\eta_i\eta_k' + F_{y_i'y_k'}\eta_i'\eta_k',$$

$$2\gamma = \left[(F_x - y_i'F_{y_i})dx + 2F_{y_i}dy_idx\right]_1^2 + 2q + 2e_\mu q_\mu$$

with dx, dy_i replaced by ξ, $y_i'\xi+\eta_i$ [12, p. 71]. The equations of variation along E are the equations

$$\text{(2:3)} \qquad \Phi_\beta(x, \eta, \eta') = 0, \qquad \Psi_\mu[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] = 0$$

in which

$$\Phi_\beta = \phi_{\beta y_i}\eta_i + \phi_{\beta y_i'}\eta_i',$$

and Ψ_μ is $d\psi_\mu$ with dx, dy_i replaced as above by ξ, $y_i'\xi+\eta_i$ [12, p. 14]. An *admissible set* ξ_1, ξ_2, $\eta_i(x)$ is one for which ξ_1, ξ_2 are constants and the functions $\eta_i(x)$ have on x_1x_2 the continuity properties of an admissible arc $y_i(x)$. The

second variation $J_2(\xi, \eta)$ for E is by definition positive definite if it is positive for all non-vanishing admissible sets $\xi_1, \xi_2, \eta_i(x)$ satisfying the equations (2:3).

IV′. An extremal arc E satisfies the condition IV′ if its second variation is positive definite.

The condition IV′ is applicable to an admissible arc which has no corners and satisfies conditions I and III′, since such an arc is necessarily non-singular and an extremal [12, pp. 112, 117].

The sufficiency theorem of Hestenes to be considered here is now the following one:

THEOREM 2:1. *If an admissible arc E has no corners and satisfies the conditions* I, II'_N, III′, IV′ *with a set of multipliers* $l_0=1$, $l_\beta(x)$, e_μ *then* $J(E)$ *is a strong relative minimum.*

Every admissible arc E satisfies the multiplier rule with none or a limited number of linearly independent non-vanishing sets of multipliers having $l_0=0$, It is said to have *order of abnormality equal to* q if it satisfies I with q and only q such sets $l_{0\sigma}=0$, $l_{\beta\sigma}(x)$, $e_{\mu\sigma}$ $(\sigma=1, \cdots, q)$. When $q=0$ it is said to be *normal.* A set of non-vanishing multipliers with $l_0=0$ will be called an *abnormal set of multipliers.*

For an admissible arc with order of abnormality equal to q the equation

$$[F_{\sigma y_i'}\eta_i]_1^2 + e_{\mu\sigma}\Psi_\mu = 0 \tag{2:4}$$

with $F_\sigma = l_{\beta\sigma}(x)\phi_\beta$ is for each σ an identity in the variables ξ_s, $\eta_{is}=\eta_i(x_s)$, since this is what the first equation (2:2) becomes for the multipliers $l_{0\sigma}=0$, $l_{\beta\sigma}(x)$, $e_{\mu\sigma}$ when the end values of dx, dy_i are replaced by those of ξ, $y_i'\xi+\eta_i$. The usual integration by parts applied to the sum

$$l_{\beta\sigma}(x)\Phi_\beta = F_{\sigma y_i}\eta_i + F_{\sigma y_i'}\eta_i'$$

gives the equation

$$\int_{x_1}^{x_2} l_{\beta\sigma}\Phi_\sigma dx = [F_{\sigma y_i'}\eta_i]_1^2, \tag{2:5}$$

so that for every admissible set of variations satisfying the equations $\Phi_\beta=0$ we find with the help of equations (2:4) and (2:5) the relations

$$[F_{\sigma y_i'}\eta_i]_1^2 = 0, \qquad e_{\mu\sigma}\Psi_\mu = 0. \tag{2:6}$$

The matrix of the q sets of values $e_{\mu\sigma}$ $(\sigma=1, \cdots, q)$ is necessarily of rank

q. Otherwise there would be a linear combination of these sets vanishing identically, and, according to a remark made above, the same combination of the linearly independent complete sets $l_{0\sigma}$, $l_{\beta\sigma}(x)$, $e_{\mu\sigma}$ would then also vanish identically, which is impossible. In the following paragraphs the variable ρ is understood to have as its range a subset of the numbers $\mu=1, \cdots, p$ such that the determinant $|e_{\rho\sigma}|$ is different from zero, and the variable ν will have the range complementary to that of ρ. The second equation (2:6) then shows that for an admissible set ξ_1, ξ_2, $\eta_i(x)$ the equations $\Psi_\rho=0$ are consequences of the equations $\Phi_\beta=\Psi_\nu=0$.

THEOREM 2:2. *Let E be an admissible arc without corners which satisfies the hypotheses of Theorem 2:1 with a set of multipliers $l_0=1$, $l_\beta(x)$, e_μ, and let ρ and ν be variables whose ranges are determined by the linearly independent abnormal sets of multipliers of E as described in the last paragraph. Then the arc E is normal for the modified problem of minimizing the functional $J(C)+e_\rho\psi_\rho(C)$ in the class of admissible arcs C satisfying the reduced system of equations $\phi_\beta=\psi_\nu=0$, and the arc E with the multipliers $l_0=1$, $l_\beta(x)$, e_ν satisfies the hypotheses of Theorem 2:1 for the modified problem. Furthermore if $J(E)+e_\rho\psi_\rho(E)$ is a strong relative minimum for the modified problem, then $J(E)$ is a similar minimum for the original problem.*

It is easy to see that the arc E is normal for the modified problem. For if E had for that problem a set of non-vanishing multipliers of the form $l_0=0$, $l_\beta(x)$, e_ν, the set $l_0=0$, $l_\beta(x)$, $e_\rho=0$, e_ν would be multipliers for E and the original problem, necessarily linearly expressible in terms of the q sets $l_{0\sigma}=0$, $l_{\beta\sigma}(x)$, $e_{\mu\sigma}$ $(\sigma=1, \cdots, q)$. This is, however, impossible on account of the fact that the determinant $|e_{\rho\sigma}|$ is not zero.

The arc E satisfies the hypotheses of Theorem 2:1 for the modified problem with the multipliers $l_0=1$, $l_\beta(x)$, e_ν, as one readily sees by an examination of the conditions I, II'_N, III$'$, IV$'$. For the condition IV$'$ one needs to note that on account of the second equation (2:6) the restricted system $\Phi_\beta=\Psi_\nu=0$ implies the complete system $\Phi_\beta=\Psi_\mu=0$.

Since the class of arcs in which a minimizing one is sought for the modified problem includes as a subclass those among which a minimizing arc is sought for the original problem, and since on the subclass the values of the functionals $J(C)+e_\rho\psi_\rho(C)$ and $J(C)$ are equal, the last statement of the theorem is evidently true.

The remarks made at the end of §1 are now applicable for the most part to the problem of Bolza also. As a result of Theorem 2:2 it is clear that no generality is lost by proving Theorem 2:1 for normal arcs only, and the proof for such arcs turns out to be in some respects simpler than for the abnormal

arcs included in the proof of Hestenes. A normal arc is a non-singular arc of the class in which a minimizing arc is sought in the sense that near every normal arc there are an infinity of other arcs of the class [12, pp. 49, 51]. The minimum problem near such an arc is therefore never trivial. Near an abnormal arc E, on the other hand, there may be no other arc of the class in which a minimizing one is sought, as in the case when $\psi_1(E)$ vanishes and is a strong relative minimum or maximum in the class of admissible arcs satisfying the conditions $\phi_\beta = \psi_2 = \cdots = \psi_p = 0$. In this case the minimum problem near E is trivial. The variety of types of abnormal arcs is evidently very great. Those included in the sufficiency theorems of Hestenes are of a special type closely related to normal arcs. Other important special types can doubtless be described and discussed, and it might be useful to have results of this kind. But it seems likely that a comprehensive theory would at this time be exceedingly elaborate and difficult, and perhaps impossible.

When the number of the end conditions $\psi_\mu = 0$ is equal to the number $2n+2$ of end values x_s, y_{is} $(s=1, 2)$ the problem is said to have fixed end points. An admissible arc E is by definition normal on a sub-interval $x'x''$ if its corresponding sub-arc is normal relative to the problem with fixed end points on that interval. The assumption that an arc E is normal on every sub-interval is evidently undesirable, for the same reason that it would be undesirable to assume for the problem of §1 that the determinants of order m of some particular set belonging to the matrix $\|\phi_{\beta y_i'}\|$ are all different from zero. For the problem of Mayer, which is the problem of Bolza with integrand function f identically zero, every minimizing arc is abnormal on every sub-interval, as has been pointed out by Carathéodory [6, 7] and others. No theory based upon the assumption of normality on sub-intervals can therefore be effective in this important case.

Bibliography

1. Bolza, *Vorlesungen über Variationsrechnung*, 1st edition, 1909; 2d edition, 1933.

2. Bliss, *The problem of Lagrange in the calculus of variations*, American Journal of Mathematics, vol. 52 (1930), pp. 673–744.

3. Morse, *Sufficient conditions in the problem of Lagrange with variable end conditions*, American Journal of Mathematics, vol. 53 (1931), pp. 517–546; see also Morse, *The problems of Lagrange and Mayer under general end conditions*, Proceedings of the National Academy of Sciences, vol. 16 (1930), pp. 229–233.

4. Morse, *Sufficient conditions in the problem of Lagrange with fixed end points*, Annals of Mathematics, vol. 32 (1931), pp. 567–577.

5. Bliss, *The problem of Bolza in the calculus of variations*, Annals of Mathematics, vol. 33 (1932), pp. 261–274.

6. Carathéodory, *Die Theorie der zweiten Variation beim Problem von Lagrange*, Sitzungsberichte der Bayerischen Akademie der Wissenschaften, 1932, pp. 99–114.

7. Carathéodory, *Über die Einteilung der Lagrangeschen Variationsprobleme nach Klassen*, Commentarii Mathematici Helvetici, vol. 5 (1933), pp. 1–10.

8. Graves, *On the Weierstrass condition for the problem of Bolza in the calculus of variations*, Annals of Mathematics, vol. 33 (1932), pp. 747–752.

9. Bliss and Hestenes, *Sufficient conditions for a problem of Mayer in the calculus of variations*, these Transactions, vol. 35 (1933), pp. 305–326.

10. Hestenes, *Sufficient conditions for the general problem of Mayer with variable end points*, these Transactions, vol. 35 (1933), pp. 479–490; also *Contributions to the Calculus of Variations* 1931–1932, The University of Chicago.

11. Hestenes, *Sufficient conditions for the problem of Bolza in the calculus of variations*, these Transactions, vol. 36 (1934), pp. 793–818.

12. Bliss, *The problem of Bolza*, Mimeographed lecture notes, University of Chicago, Winter Quarter, 1935.

13. Morse, *Sufficient conditions in the problem of Lagrange without assumptions of normalcy*, these Transactions, vol. 37 (1935), pp. 147–160.

14. Hestenes, *The problem of Bolza in the calculus of variations in parametric form*, American Journal of Mathematics, vol. 58 (1936), pp. 391–406.

15. Reid, *The theory of the second variation for the non-parametric problem of Bolza*, American Journal of Mathematics, vol. 57 (1935), pp. 573–586.

16. Hestenes, *On sufficient conditions in the problems of Lagrange and Bolza*, Annals of Mathematics, vol. 37 (1936), pp. 543–551.

University of Chicago,
Chicago, Ill.

A THEOREM IN FINITE PROJECTIVE GEOMETRY AND SOME APPLICATIONS TO NUMBER THEORY*

BY

JAMES SINGER

A point in a finite projective plane $PG(2, p^n)$, may be denoted by the symbol (x_1, x_2, x_3), where the coordinates x_1, x_2, x_3 are marks of a Galois field of order p^n, $GF(p^n)$. The symbol $(0, 0, 0)$ is excluded, and if k is a non-zero mark of the $GF(p^n)$, the symbols (x_1, x_2, x_3) and (kx_1, kx_2, kx_3) are to be thought of as the same point. The totality of points whose coordinates satisfy the equation $u_1x_1+u_2x_2+u_3x_3=0$, where u_1, u_2, u_3 are marks of the $GF(p^n)$, not all zero, is called a line. The plane then consists of $p^{2n}+p^n+1=q$ points and q lines; each line contains p^n+1 points.†

A finite projective plane, $PG(2, p^n)$, defined in this way is Pascalian and Desarguesian; it exists for every prime p and positive integer n, and there is only one such $PG(2, p^n)$ for a given p and n (VB, p. 247, VY, p. 151).

Let A_0 be a point of a given $PG(2, p^n)$, and let C be a collineation of the points of the plane. (A collineation is a 1–1 transformation carrying points into points and lines into lines.) Suppose C carries A_0 into A_1, A_1 into $A_2, \cdots, A_{k-1}$ into A_0; or, denoting the product $C\cdot C$ by C^2, $C\cdot C^2$ by C^3, etc., we have $C(A_0)=A_1$, $C^2(A_0)=A_2, \cdots, C^k(A_0)=A_0$. If k is the smallest positive integer for which $C^k(A_0)=A_0$, we call k the *period of C with respect to the point A_0*. If the period of a collineation C with respect to a point A_0 is q $(=p^{2n}+p^n+1)$, then the period of C with respect to any point in the plane is q, and in this case we will call C simply a collineation of period q.

We prove in the first theorem that there is always at least one collineation of period q, and from it we derive some results of interest in finite geometry and number theory.

Let

$$x^3 - a_3x^2 - b_3x - c_3 = 0 \tag{1}$$

be a primitive irreducible cubic belonging to a field $GF(p^n)$ which defines a $PG(2, p^n)$. A root λ of equation (1) can then be used as a generator of the

* Presented to the Society, October 27, 1934, under a different title; received by the editors April 22, 1937.

† These definitions are taken directly from the paper by Veblen and Bussey, *Finite projective geometries*, these Transactions, vol. 7 (1906), p. 244, referred to later as VB; and from the textbook by Veblen and Young, *Projective Geometry*, vol. 1, pp. 1–25, 201, referred to later as VY.

non-zero elements of a $GF(p^{3n})$ which contains the given field as a subfield. By means of the equation we can express any power of λ in terms of λ^2, λ, and 1 with coefficients in the $GF(p^n)$, that is,

$$\lambda^i = a_i\lambda^2 + b_i\lambda + c_i; \qquad i = 0, 1, \cdots . \tag{2}$$

Conversely, any three marks, a, b, c, not all zero, of the $GF(p^n)$ will uniquely determine a power of λ and therefore a non-zero mark of the $GF(p^{3n})$. We call a_i, b_i, c_i the *coordinates* of λ^i.

Since λ is a generator of the non-zero elements of the $GF(p^{3n})$, the first $p^{3n}-1$ powers of λ are distinct and $\lambda^0=\lambda^{p^{3n}-1}=1$. The powers of λ in the $GF(p^n)$ are

$$\lambda^{jq}; \qquad j = 0, 1, \cdots, p^n - 2;\ q = p^{2n} + p^n + 1. \tag{3}$$

Two non-zero marks, λ^u and λ^v, of the $GF(p^{3n})$ will be called *similar* if their ratio is a mark of the $GF(p^n)$, that is, if $u \equiv v \pmod{q}$. If the coordinates of a mark λ^u are a_u, b_u, c_u, the coordinates of a similar mark will be ka_u, kb_u, kc_u, since the coordinates of a mark in the $GF(p^n)$ are 0, 0, k.

Let the q distinct points of the plane defined by the given field be called

$$A_0, A_1, A_2, \cdots, A_{q-1}, \tag{4}$$

and suppose the notation so chosen that the coordinates of A_u are (a_u, b_u, c_u), $u=0, 1, \cdots, q-1$, where the a's, b's, and c's are given by (2). If k is any non-zero element of the $GF(p^n)$, then (ka_u, kb_u, kc_u) also are the coordinates of A_u. The possible choices for the coordinates of the point A_u then correspond to the coordinates of all the marks

$$\lambda^{u+jq}, \qquad j = 0, 1, \cdots, p^n - 2, \tag{5}$$

similar to the mark λ^u. A point A_u may then be identified with the class of similar marks (5).

Two similar non-zero marks of the $GF(p^{3n})$ are linearly dependent with respect to the $GF(p^n)$. Conversely, two non-zero linearly dependent marks of the $GF(p^{3n})$ are similar. A point can then be considered as the totality of (non-zero) marks of the $GF(p^{3n})$ linearly dependent with respect to the $GF(p^n)$ on a given non-zero mark of the $GF(p^n)$. In the same way, a line can be considered as the totality of (non-zero) marks of the $GF(p^{3n})$ linearly dependent with respect to the $GF(p^n)$ on a given pair of linearly independent marks of the $GF(p^{3n})$. The plane is the totality of (non-zero) marks of the $GF(p^{3n})$ linearly dependent with respect to the $GF(p^n)$ on a given set of three linearly independent marks of the $GF(p^{3n})$. Any four marks of the $GF(p^{3n})$

are linearly dependent with respect to the $GF(p^n)$; hence the plane exhausts all the non-zero marks of the $GF(p^{3n})$.

We now prove the following theorem:

THEOREM. *There is always at least one collineation of period* $q\ (=p^{2n}+p^n+1)$ *in the* $PG(2, p^n)$.

Consider the transformation given by

$$\begin{aligned} y_1 &= a_3x_1 + x_2, \\ y_2 &= b_3x_1 + x_3, \\ y_3 &= c_3x_1, \end{aligned} \tag{6}$$

which sends the point (x_1, x_2, x_3) into the point (y_1, y_2, y_3), a_3, b_3, c_3 being the coefficients in equation (1). A transformation of this type is a collineation, indeed, a projective collineation (VB, p. 253). But from (2) we have

$$\begin{aligned} a_{i+1} &= a_3a_i + b_i, \\ b_{i+1} &= b_3a_i + c_i, \qquad i = 0, 1, \cdots. \\ c_{i+1} &= c_3a_i, \end{aligned} \tag{7}$$

Hence the transformation (6) sends the mark λ^u into the mark λ^{u+1}, and therefore the collineation sends the point A_u into the point A_{u+1}, $u=0, 1, \cdots, q-2$. The point A_{q-1} is sent into the point A_0. The theorem is therefore proved.

This theorem has several immediate and interesting consequences. The points and lines of a $PG(2, p^n)$ can be exhibited as a rectangular array of q columns and p^n+1 rows; the elements of the array are the points, and the points in a column are the points of a line (VY). By means of the theorem we can show that the points and lines of the plane can be exhibited in a *regular* array; that is, one in which each row is a cyclic permutation of the first. For let the line containing the points A_0 and A_1 also contain the points $A_{d_2}, A_{d_3}, \cdots, A_{d_{p^n}}$. We write d_0 and d_1 for 0 and 1, respectively and for the sake of brevity, we denote a point A_u by its subscript u.

Consider the array

$$\begin{array}{cccccc} d_0 & d_0+1 & d_0+2 & \cdots & d_0+(q-2) & d_0+(q-1) \\ d_1 & d_1+1 & d_1+2 & \cdots & d_1+(q-2) & d_1+(q-1) \\ d_2 & d_2+1 & d_2+2 & \cdots & d_2+(q-2) & d_2+(q-1) \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ d_{p^n} & d_{p^n}+1 & d_{p^n}+2 & \cdots & d_{p^n}+(q-2) & d_{p^n}+(q-1). \end{array} \tag{8}$$

If all these integers are reduced modulo q, so that each lies in the range

$0, 1, \cdots, q-1$, each row will be a cyclic permutation of the first and each row will represent the totality of points (4). The integers in the $(i+1)$st column are equal to the corresponding ones of the ith column increased by unity. The collineation (6) will then carry the ith column into the $(i+1)$st (and the last column into the first) hence, since the integers in the first column represent the points of a line, the integers in any column will represent the points of a line.

The first two columns of the array (8) cannot be identical, for then q, the number of points in the plane, would equal p^n+1. They must then represent distinct lines and thus will have one and only one integer in common since two lines intersect in just one point. This implies that the first column can have only the one pair, d_0, d_1, of consecutive integers, modulo q. For if d_u, d_v is another pair of consecutive integers, where $1 \not\equiv d_v \equiv d_u+1 \pmod q$, the first two columns would have the integers d_1 and d_v in common. Since the first column cannot have more than one pair of consecutive integers, modulo q, no column can have more than one pair of consecutive integers, modulo q. It follows that no two columns of the array are identical. For if the $(i+1)$st and the $(j+1)$st were identical, we would have $d_0+i \equiv d_u+j$, $d_1+i \equiv d_v+j$, $d_v \not\equiv 1$, all modulo q. By subtracting the first congruence from the second, we see that d_u and d_v are consecutive. But this is impossible, hence the columns of the array (8) must represent the q distinct lines of the plane. The array is, therefore, a regular array exhibiting the points and lines of the plane.

The regular array leads to an interesting result in the theory of numbers. Consider those columns of the array (8) which contain the integer $0=d_0$; namely,

$$
\begin{array}{ccccc}
d_0 - d_0 & d_0 - d_1 & d_0 - d_2 & \cdots & d_0 - d_{p^n} \\
d_1 - d_0 & d_1 - d_1 & d_1 - d_2 & \cdots & d_1 - d_{p^n} \\
d_2 - d_0 & d_2 - d_1 & d_2 - d_2 & \cdots & d_2 - d_{p^n} \\
\cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & & \cdot\ \cdot\ \cdot \\
d_{p^n} - d_0 & d_{p^n} - d_1 & d_{p^n} - d_2 & \cdots & d_{p^n} - d_{p^n}.
\end{array}
\tag{9}
$$

These columns represent the pencil of lines on the point A_0. Hence in the square array (9) of p^n+1 rows and columns, the $p^n(p^n+1)$ integers not on the principal diagonal are all distinct, and therefore are congruent, modulo q, to the integers $1, 2, \cdots, p^{2n}+p^n$ in some order. The p^n+1 integers on the principal diagonal are all zero. We have thus proved the following theorem:

THEOREM. *A sufficient condition that there exist $m+1$ integers,*

$$d_0, d_1, \cdots, d_m, \tag{10}$$

having the property that their m^2+m differences d_i-d_j, $i\neq j$; $i, j=0, 1, \cdots, m$, are congruent, modulo m^2+m+1, to the integers

$$(11) \qquad 1, 2, \cdots, m^2+m$$

*in some order is that m be a power of a prime.**

We will call a set of integers such as (10) having the property described in the theorem, a *perfect difference set of order* $m+1$. If the integers (10) form a perfect difference set, so will the integers d_0+d, d_1+d, $\cdots$, d_m+d, for any d. Hence the integers in any column of the regular array (8) are a perfect difference set. Also, the integers in the set

$$(12) \qquad td_0, td_1, \cdots, td_m$$

will form a perfect difference set whenever t is relatively prime to m^2+m+1. This is true since the integers $t, 2t, 3t, \cdots, (m^2+m)t$, when reduced modulo m^2+m+1, will be a rearrangement of the integers (11).

If (10) is a perfect difference set and k is any integer in the set (11), the congruence $d_x-d_y\equiv k \pmod{m^2+m+1}$ has a unique solution $d_x=d_u$, $d_y=d_v$, d_u, d_v in (10). Consider now the set of integers

$$(13) \qquad a_0, a_1, \cdots, a_m$$

defined by the congruences

$$a_i \equiv d_{i+1} - d_i \pmod{m^2+m+1}, \qquad i = 0, 1, \cdots, m.$$

(the subscript $m+1$ is to be replaced by the subscript 0). It follows from the definition of the a's that if $k\equiv d_u-d_v$, then $k\equiv a_v+a_{v+1}+\cdots+a_{v+(u-v-1)}$, modulo m^2+m+1. That is, any integer k of (11) is congruent, modulo m^2+m+1, to a circular sum of the integers of (13), where by a circular sum we mean a sum of consecutive integers of (13), considering a_m and a_0 as consecutive. Since there are m^2+m+1 such circular sums, including the sum $a_0+a_1+\cdots+a_m$, which is congruent to 0, modulo m^2+m+1, any integer of the series

$$(14) \qquad 0, 1, 2, \cdots, m^2+m$$

is congruent to one and only one circular sum of the integers of (13). The set of integers (13) is therefore a perfect partition of m^2+m+1 in the sense of Kirkman.† It is to be noted that the order in which the integers of (13) are

* In connection with this theorem, see the proposed problem and discussion by O. Veblen, F. H. Safford, and L. E. Dickson in the American Mathematical Monthly, vol. 13 (1906), pp. 46 and 215, and vol. 14 (1907), p. 107.

† Kirkman, *On the perfect r-partitions of r^2-r+1*, Transactions of the Historical Society of Lancashire and Cheshire, vol. 9 (1857), pp. 127–142. The r of Kirkman's paper is equal to $m+1$ here. The problem of perfect partitions has been studied by a number of authors since Kirkman's time.

written is important, the same integers in a different order will usually not form a perfect partition.

If we start with the perfect partition (13), we can obtain in an obvious way the perfect difference set (10). A perfect difference set can be developed into an array such as (8). If the integers are now interpreted as points and the columns as lines, it is an easy matter to verify that the array represents the points and lines of a finite projective geometry. Whether m must be a power of a prime, and whether, if it is, the plane is necessarily Pascalian and Desarguesian, are still open questions.

Let $d_0', d_1', \cdots, d_m'$, be a perfect difference set. It will contain just one pair of consecutive integers, modulo m^2+m+1, for the congruence $d_x'-d_y'\equiv 1$ (mod m^2+m+1) has a unique solution. Suppose that $d_v'-d_u'\equiv 1$; then the set

$$d_0'-d_u', d_1'-d_u', \cdots, d_m'-d_u'$$

will be a perfect difference set and will contain the integers 0 and 1. Suppose also that each integer is reduced so that it lies in the range (14). We call such a set a *reduced* perfect difference set. Any perfect difference set leads to a unique reduced perfect difference set. Two reduced perfect difference sets will be called *identical* if they contain the same integers. The order in which these integers are written is, of course, immaterial. Two perfect difference sets will be called *equivalent* if their reduced perfect difference sets are identical. Two perfect partitions will be called *equivalent* if their corresponding perfect difference sets are equivalent. If the integers (10) of a reduced perfect difference set are written in normal order, that is if $d_i<d_{i+1}$, the corresponding perfect partition will be called *normal*. If two perfect difference sets or two perfect partitions are equivalent, the corresponding normal perfect partitions will be identical, not only with respect to the integers involved, but also with respect to the order in which they are written. Thus, any two columns of the array (8) will lead to identical normal perfect partitions, and conversely.

We now investigate the number of distinct perfect difference sets or, what is the same thing, the number of distinct perfect partitions of a given order. All known examples arise from a regular array exhibiting the points and lines of a $PG(2, p^n)$ defined by means of a $GF(p^n)$. We limit ourselves to to such perfect difference sets. The number m^2+m+1 is now $q=p^{2n}+p^n+1$.

First of all, the sets (10) and (12) are equivalent if t is a power of p. (Clearly, any power of p is relatively prime to q.) To see this, let

$$A_{d_0}, A_{d_1}, \cdots, A_{d_{p^n}},$$

be the points of the plane corresponding to the integers (10), and let

$\lambda^{d_0}, \lambda^{d_1}, \cdots, \lambda^{d_{p^n}}$ be the marks of the $GF(p^{3n})$ whose coordinates are the same as those of the points. The marks $\lambda^{td_0}, \lambda^{td_1}, \cdots, \lambda^{td_{p^n}}$ will then correspond to the integers (12). If u, v, and w are any three integers of (10), then, since A_u, A_v, and A_w are collinear, there will exist three marks k_1, k_2, k_3 of the $GF(p^n)$ such that

$$k_1\lambda^u + k_2\lambda^v + k_3\lambda^w = 0. \tag{15}$$

Raising each side of (15) to the pth power, we get

$$k_1^p\lambda^{pu} + k_2^p\lambda^{pv} + k_3^p\lambda^{pw} = 0. \tag{16}$$

(The other terms in the multinomial expansion will drop out because each coefficient will be a multiple of p and $p\equiv 0$ in the $GF(p^n)$.) Since k_1^p, k_2^p, k_3^p are in the $GF(p^n)$, equation (16) shows that the marks λ^{pu}, λ^{pv}, and λ^{pw} are linearly dependent with respect to the $GF(p^n)$. Hence the points A_{pu}, A_{pv}, and A_{pw} are collinear, and the perfect difference sets (10) and (12) are equivalent when $t=p$. The same argument shows that (10) and (12) are equivalent when t is equal to any power of p.

Secondly, it appears from all known examples, although a general proof is still lacking, that (10) and (12) will be distinct if t is prime to q and is not a power of p (mod q). However, if $t=-1$, the sets will be distinct since in this case the integers in the normal perfect partition corresponding to the perfect difference set (12) will be the same as those in the normal perfect partition corresponding to the perfect difference set (10), but in reverse order. Since a perfect partition cannot contain two equal integers, a perfect partition is necessarily distinct from its inverse; hence the set (10) will be distinct from its inverse (12) for $t=-1$.

It also seems to be true that if (10) and

$$d_0', d_1', \cdots, d_m' \tag{17}$$

are any two perfect difference sets of the same order, there is a t for which (12) and (17) are equivalent. If these statements are true, the number of distinct perfect difference sets (or the number of distinct perfect partitions) for a given p^n is equal to

$$\frac{\phi(q)}{3n},$$

where $\phi(q)$ is the Euler function, the number of positive integers not greater than and prime to q. This number is even, since each perfect difference set can be paired with its inverse.

I append a partial list of the (reduced) perfect difference sets and their

corresponding normal perfect partitions. I give a single set for each p^n, the remaining ones can easily be found by the methods given above.

p^n	q	$\phi(q)/3n$	perfect difference set	perfect partition
2	7	2	0 1 3	1 2 4
2^2	21	2	0 1 4 14 16	1 3 10 2 5
2^3	73	8	0 1 3 7 15 31 36 54 63	1 2 4 8 16 5 18 9 10
2^4	273	12	0 1 3 7 15 31 63 90 116 127 136 181 194 204 233 238 255	1 2 4 8 16 32 27 26 11 9 45 13 10 29 5 17 18
3	13	4	0 1 3 9	1 2 6 4
3^2	91	12	0 1 3 9 27 49 56 61 77 81	1 2 6 18 22 7 5 16 4 10
5	31	10	0 1 3 8 12 18	1 2 5 4 6 13
7	57	12	0 1 3 13 32 36 43 52	1 2 10 19 4 7 9 5
11	133	36	0 1 3 12 20 34 38 81 88 94 104 109	1 2 9 8 14 4 43 7 6 10 5 24
13	183	40	0 1 3 16 23 28 42 76 82 86 119 137 154 175	1 2 13 7 5 14 34 6 4 33 18 17 21 8

The preceding concepts are susceptible of immediate generalization. Let

$$x^{k+1} - a_{k+1,1}x^k - a_{k+1,2}x^{k-1} - \cdots - a_{k+1,k+1} = 0 \tag{1'}$$

be a primitive irreducible $(k+1)$st degree equation belonging to a $GF(p^n)$. A root λ of the equation is a generator of the non-zero elements of a $GF(p^{(k+1)n})$. By means of the equation, we can express any power of λ in terms of λ^k, $\lambda^{k-1}, \cdots, 1$, that is,

$$\lambda^i = a_{i,1}\lambda^k + a_{i,2}\lambda^{k-1} + \cdots + a_{i,k+1}, \qquad i = 0, 1, \cdots. \tag{2'}$$

Conversely, any $k+1$ marks $a_1, a_2, \cdots, a_{k+1}$, not all zero, of the $GF(p^n)$ will uniquely determine a power of λ, and hence a non-zero mark of the $GF(p^{(k+1)n})$. The $k+1$ marks will be called the coordinates of that power of λ.

The $GF(p^{(k+1)n})$ defines a k-dimensional finite projective geometry, $PG(k, p^n)$. An h-dimensional space, $h=0, 1, \cdots, k$, is defined as the totality of marks of the $GF(p^{(k+1)n})$ linearly dependent with respect to the included $GF(p^n)$ on $h+1$ linearly independent marks of the $GF(p^{(k+1)n})$. A point is a zero-space, a line is a 1-space, etc. Any $h+1$ linearly independent marks of an h-space will define the same h-space. These definitions are equivalent to those in the paper by Veblen and Bussey (loc. cit.) if the coordinates of a point are interpreted as the coordinates of any mark in a class of similar marks.

Let the sum $p^{hn}+p^{(h-1)n}+\cdots+p^n+1$ be denoted by q_h, $h=0, 1, \cdots$. If the q_k distinct points of the $PG(k, p^n)$ are denoted by the integers

$$0, 1, \cdots, q_k - 1, \tag{4'}$$

the points, lines, planes, etc., of the geometry can be exhibited as a regular array in the form of a k-dimensional rectangular matrix whose elements are these integers. The integers in a properly chosen $(k-1)$-dimensional face of the matrix represent the points of a $(k-1)$-space. The remaining $(k-1)$-spaces are the $(k-1)$-dimensional layers parallel to this face. The integers in these layers are obtained by successively adding 1's to the integers of the first face. The integers in a properly chosen $(k-2)$-dimensional face of a $(k-1)$-dimensional face or layer represent the points of a $(k-2)$-dimensional space, etc. The existence of this regular array follows from the existence of the transformation

$$\begin{aligned} y_1 &= a_{k+1,1}x_1 + x_2, \\ y_2 &= a_{k+1,2}x_1 + x_3, \\ &\cdots\cdots\cdots \\ y_k &= a_{k+1,k}x_1 + x_{k+1}, \\ y_{k+1} &= a_{k+1,k+1}x_1. \end{aligned} \tag{6'}$$

This transformation sends an h-space into an h-space since it preserves linear dependence. It sends the mark λ^u into the mark λ^{u+1}. The regular array can then be constructed.

The regular array yields the difference set of q_{k-1} integers

$$d_0, d_1, \cdots, d_{q_{k-1}-1} \tag{10'}$$

having the property that their differences, d_i-d_j, $i\neq j$; $i, j=0, 1, \cdots, q_{k-1}-1$, are congruent, modulo q_k, to the integers

$$1, 2, \cdots, q_k - 1, \tag{11'}$$

each integer of (11′) being congruent to q_{k-2} of the differences. The difference set (10′) leads to a partition

$$a_0, a_1, \cdots, a_{q_{k-1}-1} \tag{13'}$$

having the property that each of the integers in (11′) is congruent, modulo q_k, to exactly q_{k-2} circular sums of (13′). The sum of all the integers of (13′) is congruent to 0, modulo q_k.

BROOKLYN COLLEGE,
BROOKLYN, N. Y.

SYMMETRIC AND ALTERNATE MATRICES IN AN ARBITRARY FIELD, I*

BY

A. ADRIAN ALBERT

Introduction

The elementary theorems of the classical treatment of symmetric and alternate matrices may be shown, without change in the proofs, to hold for matrices whose elements are in any field of characteristic not two. The proofs fail in the characteristic two case and the results cannot hold since here the concepts of symmetric and alternate matrices coincide. But it is possible to obtain a unified treatment. We shall provide this here by adding a condition to the definition of alternate matrices which is redundant except for fields of characteristic two. The proofs of the classical results will then be completed by the addition of two necessary new arguments.

The theorems on the definiteness of real symmetric matrices have had no analogues for general fields. They have been based on the property that the sum of any two non-negative real numbers is non-negative. This is equivalent to the property that for every real a and b we have $a^2+b^2=c^2$ for a real c. But $a^2+b^2=(a+b)^2$ in any field of characteristic two and we shall use this fact to obtain complete analogues for arbitrary fields of characteristic two of the usual theorems on the definiteness of real symmetric matrices.

Quadratics forms may be associated with symmetric matrices and the problem of their equivalence is equivalent to the problem of the congruence of the corresponding matrices. This is true except when the field of reference has characteristic two where no matric treatment has been given. We shall associate quadratic forms in this case with a certain type of non-symmetric matrix and shall use our results on the congruence of alternate matrices to obtain a matrix treatment of the quadratic form problem.

The classical theorems† on pairs of symmetric or alternate matrices with complex elements will be shown here to be true for matrices with elements in any algebraically closed field whose characteristic is not two. This will be seen to imply that any two symmetric (or alternate) matrices are orthogonally equivalent if and only if they are similar. But the proof fails for fields of characteristic two.

* Presented to the Society, April 10, 1937; received by the editors April 26, 1937.

† Cf. L. E. Dickson, *Modern Algebraic Theories*, chap. 6. See also the report of C. C. MacDuffee, Ergebnisse der Mathematik, vol. 21 (1933), part 5, for this material as well as the classical results referred to above. The theory will also be found in J. H. M. Wedderburn's *Lectures on Matrices*, American Mathematical Society Colloquium Publications, vol. 17, 1934.

We shall prove the existence of two similar symmetric matrices with elements in a field $\mathfrak{F}$ of characteristic two which are not orthogonally equivalent in the algebraically closed extension of $\mathfrak{F}$. Our treatment of the theory of the orthogonal equivalence in $\mathfrak{F}$ of characteristic two will be rational, that is, no algebraic closure properties of $\mathfrak{F}$ will be assumed. Our formulation will involve a recasting of the theory of similarity of square matrices and then a corresponding parallel treatment of the theory of orthogonal equivalence. In particular we shall obtain a complete determination of the invariant factors of any symmetric matrix in $\mathfrak{F}$ of characteristic two.

The generalized transposition concept called an involution* J of the set of all n-rowed square matrices arises naturally in any rational treatment of orthogonal equivalence. The consequent study of the J-orthogonal equivalence of J-symmetric and J-alternate matrices will be introduced here and various important special types treated in subsequent papers.

I. Congruence theory

1. **Elementary concepts.** Let $\mathfrak{F}$ be an arbitrary field, and let $A=(a_{ij})$ $(i,j=1,\cdots,n)$ be an n-rowed square matrix with elements a_{ij} in $\mathfrak{F}$. We use the customary notation A' for the transpose of A and call A symmetric if $A'=A$. We shall modify the usual definition of alternate matrices however, and make the classically consistent definition:

Definition. *A matrix A is called alternate (or skew-symmetric) if $A'=-A$, and the diagonal elements a_{ii} of A are all zero.*

Notice that when the characteristic of $\mathfrak{F}$ is not two the final part of our definition is redundant. But when the characteristic is two every symmetric matrix has the property $A=-A'$ and the condition will be shown to be essential. We shall also call a matrix A *non-alternate symmetric* if $A=A'$ and A is not alternate according to our definition above. This last condition is redundant except for fields of characteristic two, in which case we are simply assuming that $A=A'$ has a non-zero diagonal element.

Two square matrices A and B with elements in a field $\mathfrak{F}$ are said to be congruent in $\mathfrak{F}$ if there exists a non-singular matrix P with elements in $\mathfrak{F}$ such that $B=PAP'$. It is easy to prove the following lemmas:†

Lemma 1. *Let B be obtained from A by any permutation of its rows followed by the same permutation on the columns. Then B and A are congruent in $\mathfrak{F}$.*

* See the author's paper, *Involutorial simple algebras and real Riemann matrices*, Annals of Mathematics, vol. 36 (1935), pp. 886–964, p. 894 for the definition and some elementary properties of involutions.

† The proofs of these lemmas may be found in the author's *Modern Higher Algebra*, chap. 5, University of Chicago Press, 1937.

LEMMA 2. *Replace the ith row of A by the sum of this row and any linear combination of the remaining rows. Follow this with the corresponding column replacement. Then the resulting matrix B is congruent to A.*

If G and H are square matrices, the matrix

$$A = \begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} \tag{1}$$

is called the direct sum of G and H. This notion has an immediate generalization to the direct sum

$$\begin{pmatrix} G_1 & & \\ & \ddots & \\ & & G_s \end{pmatrix} \tag{2}$$

of square matrices G_i* called the components of A. It is clear that A is symmetric if and only if its components are symmetric. Also A is alternate if and only if the components of A are all alternate. But in a field of characteristic two a matrix A may be non-alternate symmetric and yet may have some alternate components but has at least one non-alternate component. We shall show later that *such matrices are always congruent to diagonal matrices*, that is, *to direct sums of one-rowed square matrices*. Our proofs will depend partly on the almost trivial consequence of Lemmas 1 and 2.

LEMMA 3. *Let A have the form* (1). *Then A is congruent to*

$$\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix},$$

for any G_0 congruent to G and H_0 congruent to H.

A principal sub-matrix of A is a sub-matrix whose main diagonal is a part of the main diagonal of A. The determinants of principal sub-matrices are called principal minors of A. Then it may easily be shown† that Lemmas 1 and 2 yield the following theorem:

LEMMA 4. *Let G be a non-singular principal sub-matrix of a symmetric or alternate matrix A. Then A is congruent in $\mathfrak{F}$ to*

$$B = \begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}$$

whose principal minors having $|G|$ as sub-determinant have the same values as those of A.

* We shall henceforth use the notation diag$[G_1, \cdots, G_s]$ for (2) to simplify printing.

† See the author's *Modern Higher Algebra*, chap. 5.

The one-rowed principal minors of A are the elements a_{ii}. When they are all zero and $A' = \pm A$ the two-rowed principal minors are

$$\begin{vmatrix} 0 & a_{ij} \\ a_{ji} & 0 \end{vmatrix} = \pm (a_{ij})^2 = 0 \qquad (i, j = 1, \cdots, n), \tag{3}$$

if and only if $A = 0$. Thus we have the following lemma:

LEMMA 5. *Every symmetric or alternate matrix $A \neq 0$ has a non-zero one- or two-rowed principal minor.*

We shall close our discussion of the tools of our theory by proving the lemma:

LEMMA 6. *Let A and B be r-rowed non-singular matrices, and let*

$$A_0 = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \qquad B_0 = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \tag{4}$$

be n-rowed square matrices. Then $B_0 = PA_0P'$ for a non-singular P if and only if

$$P = \begin{pmatrix} Q & S \\ 0 & R \end{pmatrix}, \qquad QAQ' = B, \tag{5}$$

where Q and R are non-singular.

It is clear that the lemma implies that if $QAQ' = B$ for Q non-singular, we may choose any R and S such that R is non-singular and obtain $PA_0P' = B_0$. These results are an immediate consequence of the computation

$$PA_0P' = \begin{pmatrix} Q & S \\ K & R \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} Q' & K' \\ S' & R' \end{pmatrix} = \begin{pmatrix} QAQ' & QAK' \\ KAQ' & KAK' \end{pmatrix} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \tag{6}$$

if and only if $QAQ' = B$, $QAK' = KAQ' = KAK' = 0$. But B is non-singular and so is A. Hence Q is non-singular, so therefore is AQ', and $K = 0$ is our only condition.

2. **Congruence of alternate matrices.** We shall prove the following theorem:

THEOREM 1. *Every matrix congruent to an alternate matrix is an alternate matrix.*

For let $x_1, \cdots, x_n$ be independent indeterminates over $\mathfrak{F}$

$$x = (x_1, \cdots, x_n), \qquad A = (a_{ij}) \qquad (i, j = 1, \cdots, n). \tag{7}$$

If $a_{ji} = -a_{ij}$, the quadratic form

$$xAx' = \sum_{i,j} x_i a_{ij} x_j = \sum_{i=1}^{n} a_{ii} x_i^2 . \tag{8}$$

When also A is alternate the a_{ii} are all zero and

$$xAx' \equiv 0.$$

We let $B = PAP'$, $y = (y_1, \cdots, y_n)$, and have $B' = -B$. But

$$yBy' = \sum b_{ii}y_i^2 = xAx', \qquad x = yP,$$

so that $yBy' = xAx' \equiv 0$ in the y_i and the $b_{ii} = 0$. Hence B is alternate.

The proof is of course unnecessary for fields $\mathfrak{F}$ of characteristic not two. Notice that it implies the theorem:

THEOREM 2. *Every matrix congruent to a non-alternate symmetric matrix is non-alternate symmetric.*

We may also easily prove the following theorem:

THEOREM 3. *Two alternate matrices are congruent in $\mathfrak{F}$ if and only if they have the same rank $2t$.*

For if $A \neq 0$ is alternate, its diagonal elements are all zero. By Lemma 5 A has a two-rowed principal minor

$$E_0 = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \tag{9}$$

with $a \neq 0$. By Lemma 4 the matrix A is congruent to

$$\begin{pmatrix} E_0 & 0 \\ 0 & A_0 \end{pmatrix}.$$

But

$$E = \begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}\begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ a & 0 \end{pmatrix}\begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{10}$$

so that by Lemma 3, A is congruent to

$$\begin{pmatrix} E & 0 \\ 0 & A_0 \end{pmatrix}.$$

We apply Theorem 1 to see that A_0 is alternate. A repetition of this process by the use of Lemma 3 shows that A is congruent to the direct sum

$$\operatorname{diag}[E_1, \cdots, E_t, 0], \qquad E_i = E, \tag{11}$$

where $2t$ is the rank of A. Any other alternate matrix of rank $2t$ is congruent to (11) and hence to A, and we have Theorem 3.

The proof given above is valid for general fields only because we have proved Theorem 1. Notice that we have the following consequence:

THEOREM 4. *Every alternate matrix of rank 2t is congruent in $\mathfrak{F}$ to*

$$\begin{bmatrix} 0 & -I_t & 0 \\ I_t & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{12}$$

where I_t is the t-rowed identity matrix.

This new form of Theorem 3 is a consequence of (11) and Lemma 1.

3. **Congruence of non-alternate symmetric matrices.** The Lemmas 1, 3, 4, 5 may be applied to an arbitrary symmetric matrix. They show that every symmetric matrix is congruent to a direct sum of matrices of the forms

$$(a), \qquad \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}, \qquad (a \neq 0 \text{ in } \mathfrak{F}).$$

But as in (10) we have

$$P = \begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix}, \qquad P\begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}P^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{13}$$

By Lemma 1 we have the preliminary reduction given by the following lemma:

LEMMA 7. *Every symmetric matrix is congruent in $\mathfrak{F}$ to a matrix*

$$\operatorname{diag}\,[D, G, 0], \tag{14}$$

where D is a diagonal matrix with elements in $\mathfrak{F}$ and G is a direct sum of two-rowed matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{15}$$

The reduction of symmetric matrices with elements in a field $\mathfrak{F}$ of characteristic not two is evidently completed by the fact that

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{16}$$
$$= \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix},$$

and then that the matrix G is congruent to a diagonal matrix whose diagonal elements are all 2 or -2. But the corresponding transformation in $\mathfrak{F}$ of characteristic two is clearly singular. We complete our reduction in this case by the computation in the following theorem:

THEOREM 5. *Let $a \neq 0$ be in $\mathfrak{F}$ of characteristic two and*

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & a \\ 1 & 1 & a \end{pmatrix}. \tag{17}$$

Then

$$PAP' = aI_3,$$

where I_3 is the three-rowed identity matrix.

The result above seems quite remarkable, as one might expect that the matrix A which has an alternate sub-matrix would not be congruent to a multiple of the identity matrix. We verify the computation in

$$\begin{aligned} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & a \\ 1 & 1 & a \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} P' &= \begin{pmatrix} a & 0 & 1 \\ a & a & 0 \\ a & a & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & a & a \end{pmatrix} \\ &= \begin{pmatrix} a & a+a & a+a \\ a+a & a & a+a \\ a+a & a+a & a+a+a \end{pmatrix} = aI_3, \end{aligned} \tag{18}$$

since $a+a=0$.

As an immediate consequence of Theorem 5 and Lemma 7 we have the following theorem:

THEOREM 6. *Every non-alternate symmetric matrix is congruent to a diagonal matrix.*

The problem of finding when two diagonal matrices are congruent is not solvable in a general field, as the structural properties of the field are involved in this question. It is usual then to assume that $\mathfrak{F}$ is a field such that for every a of $\mathfrak{F}$ there exists a b in $\mathfrak{F}$ such that $b^2=a$. Then for this case *two non-alternate symmetric matrices are congruent if and only if they have the same rank*. Finally, as in the classical theory, we may obtain a so-called Kronecker reduction of non-alternate symmetric matrices to diagonal form. The results are nearly the same as in the classical theory; they depend essentially on Lemma 6, and we shall not give the proofs. The only difference is that when in the Kronecker reduction we obtain a matrix

$$\begin{pmatrix} A_{r-2} & 0 \\ 0 & E \end{pmatrix},$$

with E the matrix of (15), we use (16) if $\mathfrak{F}$ does not have characteristic two

and obtain a corresponding pair of diagonal elements 2, -2. But when $\mathfrak{F}$ has characteristic two the Kronecker reduction is completed by the use of Theorem 5. Thus we replace E by

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix},$$

where a is any diagonal element obtained at any stage of the reduction.

4. **Definite symmetric matrices.** The field $\mathfrak{R}'$ of all real numbers has the characteristic property that if a and b are in $\mathfrak{R}'$, there exists a c in $\mathfrak{R}'$ such that $c^2 = a^2 + b^2$. This result has the analogue $a^2 + b^2 = (a+b)^2$ in fields of characteristic two, and we shall use these results to obtain an analogue of the concept of definiteness of real symmetric matrices.

DEFINITION. *A symmetric matrix A with elements in a field $\mathfrak{F}$ will be called semi-definite if A is congruent in $\mathfrak{F}$ to*

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

where r is the rank of A, and I_r is the r-rowed identity matrix. A non-singular semi-definite matrix will be called definite.

In the remainder of the section we assume that $\mathfrak{F}$ has characteristic two.* Clearly Theorem 1 implies that alternate matrices are never semi-definite. However by Theorems 4 and 5 the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \tag{19}$$

is semi-definite for every alternate matrix A. We use this result in the proof of the following theorem:

THEOREM 7. *A non-alternate symmetric matrix with elements in $\mathfrak{F}$ of characteristic two is semi-definite if and only if its diagonal elements are the squares of elements of $\mathfrak{F}$.*

If $A \neq 0$ is one-rowed, it has the form (a^2), is congruent to I_1, and is definite. Hence our theorem is true for one-rowed matrices, and we make an induction on the order of A. Let $A = (a_{ij})$ be n-rowed,

$$a_{ij} = a_{ji}, \qquad a_{ii} = a_i^2 \qquad\qquad (a_i, a_{ij} \text{ in } \mathfrak{F}).$$

At least one $a_i \neq 0$, and there is no loss of generality if we assume that $a_1 \neq 0$.

* As the field of reference in what follows will sometimes be general and sometimes of characteristic two we shall henceforth designate that it has characteristic two by writing $\mathfrak{F}^{(2)}$ except when the condition is explicitly stated.

Multiply the first row and column of A by a_1^{-1} and replace a_{11} by 1, A by a congruent matrix $B=(b_{ij})$, $b_{ii}=b_i^2$, $b_{ij}=b_{ji}$. We add b_{i1} times the first row of B to its ith row and replace b_{i1} by 0, b_{ii} by $b_{ii}+b_{i1}^2=(b_i+b_{i1})^2=c_i^2$. The corresponding column transformation then replaces b_{1i} by 0, and leaves c_i^2 unaltered. Thus A is congruent to the direct sum

$$\begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix}, \tag{20}$$

where the diagonal elements of C are c_i^2, c_i in $\mathfrak{F}^{(2)}$. If C is alternate we have seen that the matrix (20) is semi-definite. Otherwise C is semi-definite by our induction and so is A.

Conversely let A be semi-definite so that

$$A = \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P_1' & P_3' \\ P_2' & P_4' \end{pmatrix} = \begin{pmatrix} P_1P_1' & P_1P_3' \\ P_3P_1' & P_3P_3' \end{pmatrix}.$$

Then if $P_1=(d_{ij})$, the diagonal elements of P_1P_1' have the form $\sum_{j=1}^r d_{ij}^2 = (\sum_{j=1}^r d_{ij})^2$. Similarly the diagonal elements of P_3P_3' are squares.

COROLLARY. *The principal minors of a semi-definite symmetric matrix are the squares of elements of* $\mathfrak{F}^{(2)}$.

For every principal sub-matrix B of a semi-definite matrix is semi-definite by our theorem. If $|B|=0$, our result is true. Let $|B|\neq 0$, so that B is definite and $B=PP'$, $|B|=|P|^2$ as desired.

The classical result on real symmetric matrices states that A is positive semi-definite if and only if every principal minor of A is non-negative, that is, the square of a real number. Theorem 7 has a weaker hypothesis than the theorem about the real field but, in view of our corollary, the same conclusion. Thus our result is a true analogue of the corresponding real theorem.

In a later section we shall require the theorem:

THEOREM 8. *Let A be a semi-definite matrix with elements in an infinite field $\mathfrak{F}$ of characteristic two. Then there exist quantities a in $\mathfrak{F}$ such that*

$$a^4I + A \tag{21}$$

is definite.

For the diagonal elements of (21) are squares in $\mathfrak{F}^{(2)}$ when this is true of A. The determinant $d(a)$ of (21) is a polynomial in a with leading coefficient unity, and thus there exist infinitely many elements a in $\mathfrak{F}^{(2)}$ such that $d(a)\neq 0$, (21) is definite.

5. **Hermitian matrices.** The classical theory of the conjunctivity of Hermitian matrices with elements in a field $\mathfrak{K}$ holds for arbitrary fields. To verify

this note that the theory has already been shown to be valid for fields of characteristic not two.* Let now $\mathfrak{F}^{(2)}$ be a field of characteristic two, and let $\mathfrak{K}$ be a separable quadratic field over $\mathfrak{F}^{(2)}$. This is the only case that need be considered. Then

$$\mathfrak{K} = \mathfrak{F}^{(2)}(\theta), \qquad \theta^2 = \theta + c, \tag{22}$$

so that every k of $\mathfrak{K}$ has the form

$$k = k_1 + k_2\theta \qquad (k_1,\ k_2 \text{ in } \mathfrak{F}^{(2)}). \tag{23}$$

The correspondence

$$k \longleftrightarrow \bar{k} = k_1 + k_2(\theta + 1) \tag{24}$$

is an automorphism of $\mathfrak{K}$ over $\mathfrak{F}^{(2)}$ with the property $\bar{\bar{k}} = k$. The theory of Hermitian matrices is then a theory of matrices A with elements in $\mathfrak{K}$. Write

$$A = (a_{ij}), \qquad \bar{A} = (\bar{a}_{ij}). \tag{25}$$

Then $(\bar{A}') = \bar{A}'$ is called the conjugate transpose of A; and we call a matrix A Hermitian if $A = \bar{A}'$.

Two Hermitian matrices A and B with elements in $\mathfrak{K}$ are said to be conjunctive in $\mathfrak{K}$ if

$$B = DA\bar{D}' \tag{26}$$

for a non-singular D with elements in $\mathfrak{K}$. It is clear that all of the results leading up to our reduction theorem to diagonal form hold.

Our reduction theory is now an immediate consequence of

$$\begin{pmatrix} 1 & \theta \\ 1 & \bar{\theta} \end{pmatrix} E \begin{pmatrix} 1 & 1 \\ \bar{\theta} & \theta \end{pmatrix} = \begin{pmatrix} \theta & 1 \\ \bar{\theta} & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \bar{\theta} & \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{27}$$

since $\bar{\theta} = \theta + 1$, $\theta + \bar{\theta} = 1$. We combine this result with the Hermitian analogue of Lemma 7 and have proved the following theorem:

THEOREM 9. *Every Hermitian matrix is conjunctive in $\mathfrak{K}$ to a diagonal matrix with elements in $\mathfrak{F}^{(2)}$.*

When $\mathfrak{F}$ is a perfect field† we have the usual result:

THEOREM 10. *Any two Hermitian matrices with elements in $\mathfrak{K}$ over a perfect $\mathfrak{F}$ of characteristic two are conjunctive in $\mathfrak{K}$ if and only if they have the same rank.*

* Cf. the author's *Modern Higher Algebra*. The theory is almost exactly the same as in L. E. Dickson's *Modern Algebraic Theories*.

† A perfect field $\mathfrak{F}$ of characteristic two has the property that every a of $\mathfrak{F}$ is equal to b^2, b in $\mathfrak{F}$. Such fields with an over-field $\mathfrak{K} = \mathfrak{F}(\theta)$ exist. For the definition see van der Waerden's *Moderne Algebra*, vol. 1, as well as the author's own *Modern Higher Algebra*.

The proof of the above result is trivial and will be omitted.

II. Quadratic forms

1. The matrices of a quadratic form. Let

$$x = (x_1, \cdots, x_n), \qquad y = (y_1, \cdots, y_n)$$

with independent indeterminates $x_1, \cdots, y_n$ over any field $\mathfrak{F}$. If A is a symmetric matrix with elements in $\mathfrak{F}$ the form xAy is called a symmetric bilinear form in the variables x_i, y_j. A trivial computation then shows that two such bilinear forms are equivalent if and only if their matrices are congruent. Thus the symmetric matrix and the symmetric bilinear form theories are equivalent. Analogous results evidently hold when A is Hermitian, $xA\bar{y}$ is an Hermitian bilinear form. Moreover the theory of Hermitian quadratic forms $xA\bar{x}$ is also equivalent to the theory of Hermitian matrices since we may always choose $x_1, \cdots, x_n, \bar{x}_1, \cdots, \bar{x}_n$ to be independent indeterminates over $\mathfrak{K}$. In the theory of fields of characteristic not two the theory of quadratic forms is equivalent to that of symmetric matrices. This is not true for fields $\mathfrak{F}$ of characteristic two, and we shall develop this theory here. It was developed for the case of a finite field by L. E. Dickson (American Journal of Mathematics, vol. 21 (1899), p. 194), but the results obtained there do not hold for an arbitrary field $\mathfrak{F}$ of characteristic two. We introduce the theory as follows:

Every quadratic form in n independent indeterminates has the form

$$f = f(x_1, \cdots, x_n) = \sum_{i,j}^{1,\cdots,n} x_i a_{ij} x_j. \tag{28}$$

This expression is clearly not unique, but if also

$$f = \sum_{i,j} x_i a_{0ij} x_j, \tag{29}$$

then, by equating the coefficients of x_i^2 and $x_i x_j$, we have

$$a_{ii} = a_{0ii}, \qquad a_{ij} + a_{ji} = a_{0ij} + a_{0ji}. \tag{30}$$

Write $x = (x_1, \cdots, x_n)$, $A = (a_{ij})$, $A_0 = (a_{0ij})$, so that

$$f = xAx' = xA_0x'. \tag{31}$$

Then (30) states that $A + A' = A_0 + A_0'$, A and A_0 have the same diagonal elements. The matrix $A_0 - A$ has zero diagonal elements and $(A_0 - A)' = A_0' - A' = A - A_0 = -(A_0 - A)$. Conversely let $f = xAx'$ and $A_0 = A + N$, where $N = -N'$ and the diagonal elements of N are all zero. Then $A_0 + A_0' = A + A'$, $f = xA_0x'$. We have proved the following theorem:

THEOREM 11. *Let $f=xAx'$, $g=xA_0x'$, $A_0=A+N$. Then $f=g$ if and only if N is an alternate matrix.*

In the study of quadratic forms with coefficients in a field $\mathfrak{F}$ of characteristic not two it is customary to choose a unique matrix

$$A = (a_{ij}), \qquad a_{ij} = \tfrac{1}{2}(a_{0ij} + a_{0ji}) \tag{32}$$

so that A is symmetric. Then the theory of quadratic forms is equivalent to that of symmetric matrices. This is impossible in $\mathfrak{F}^{(2)}$ of characteristic two both because (32) is impossible and because if $A=(a_{ij})=A'$, then $f=xAx' =\sum_{i=1}^{n} a_{ii}x_i^2$. It is now natural to make the following definition:

DEFINITION. *A quadratic form f with coefficients in a field $\mathfrak{F}$ of characteristic two is called diagonal or non-diagonal according as f does or does not have the expression*

$$f = a_1x_1^2 + \cdots + a_nx_n^2 \qquad (a_i \text{ in } \mathfrak{F}). \tag{33}$$

We shall now assume that the characteristic of $\mathfrak{F}^{(2)}$ is two. By a non-singular linear transformation

$$x_i = \sum_{i=1}^{n} y_i d_{ij} \tag{34}$$

with matrix $D=(d_{ij})$ we carry a quadratic form f into what is called an equivalent form g. If A is a matrix, $f=xAx'$, and $y=(y_1, \cdots, y_n)$, then $x=yD$,

$$g = yDA(yD)' = yDAD'y'. \tag{35}$$

Hence DAD' is one possible matrix of g.

Consider in particular the case where

$$f = x_1^2 + x_2^2, \qquad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{36}$$

and write

$$x_1 = y_1 + y_2, \qquad x_2 = y_2, \qquad x = (y_1, y_2)\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = yD. \tag{37}$$

Then in a field $\mathfrak{F}^{(2)}$ we have

$$y_1 = x_1 + x_2, \qquad f = g = y_1^2, \tag{38}$$

since $(x_1+x_2)^2=x_1^2+x_2^2$. A matrix of g is

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{39}$$

which is not merely incongruent to the two-rowed identity matrix A but does not even have the same rank. However

$$DAD' = DD' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = B + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{40}$$

We now consider the general theory of quadratic forms in the light of the above example. Write

$$f = \sum_{i \leq j}^{i,j=1,\cdots,n} x_i a_{ij} x_j. \tag{41}$$

Then the matrix

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2\,n-1} & a_{2n} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & a_{n-1\,n-1} & a_{n-1\,n} \\ 0 & 0 & \cdots & 0 & a_{n\,n} \end{pmatrix} \tag{42}$$

is uniquely determined by f. We make a non-singular transformation $x = yD$ and carry f into

$$g = yBy', \tag{43}$$

where B has the form $B = (b_{ij})$, $b_{ij} = 0$ for $i > j$. Then

$$B - DAD' = N \tag{44}$$

is alternate by Theorem 11.

THEOREM 12. *Let f be a quadratic form with unique matrix A of* (42). *Then a non-singular transformation with matrix D carries f into an equivalent form with matrix $B = DAD' + N$ where N is the unique alternate matrix chosen so that the elements below the diagonal in B are all zero.*

2. **Diagonal quadratic forms.** If $A = A'$, then $B = DAD' + N$ is symmetric. Hence $g = yBy'$ is a diagonal quadratic form. This gives the following theorem:

THEOREM 13. *Every quadratic form equivalent to a diagonal quadratic form is a diagonal quadratic form in $\mathfrak{F}^{(2)}$.*

A simpler proof is given as follows. We let $f = a_1x_1^2 + \cdots + a_nx_n^2$ and use (34). Then

$$f = g(y_1, \cdots, y_n) = \sum_i a_i \Big(\sum_j d_{ij} y_j \Big)^2 = \sum_{j=1}^n \Big(\sum_{i=1}^n a_i d_{ij}^2 \Big) y_j^2 = \sum_{j=1}^n b_j y_j^2,$$

since $(a+b)^2 = a^2 + b^2$ in $\mathfrak{F}^{(2)}$. However we have now proved the theorem:

THEOREM 14. *Two diagonal quadratic forms $f=\sum_{i=1}^{n}a_ix_i^2$ and $g=\sum_{i=1}^{n}b_iy_i^2$ are equivalent in $\mathfrak{F}^{(2)}$ if and only if the coefficients b_i are representable as values $f(d_{1j},\cdots,d_{nj})$ of f such that the corresponding determinant $|d_{ij}|\neq 0$.*

We cannot go into questions of representation in a general field $\mathfrak{F}$. However we do have the following theorem:

THEOREM 15. *Let $\mathfrak{F}^{(2)}$ be perfect. Then any two non-zero diagonal quadratic forms in n indeterminates are equivalent.*

For

$$f = \sum_{i=1}^{n} a_ix_i^2 = \sum_{i=1}^{n} (\alpha_ix_i)^2 = \left(\sum_{i=1}^{n} \alpha_ix_i\right)^2 = z_1^2,$$

where $\alpha_i^2=a_i$, α_i in $\mathfrak{F}$. Similarly

$$g = \sum b_iy_i^2 = z_1^2 .$$

Then both f and g are equivalent to the same quadratic form and hence to each other.

3. **Non-diagonal quadratic forms.** Non-diagonal quadratic forms f are not equivalent to diagonal quadratic forms by Theorem 13. Then $f=xAx'$, $A+A'\neq 0$. By Theorem 12 we have $B=DAD'+N$, $N+N'=0$, and $B+B'=D(A+A')D'$. The matrix $A+A'$ is alternate and has rank $2r$ by Theorem 3. Moreover we may always carry f into a form g whose matrix B has the property

$$D(A+A')D' = B + B' = \begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad W_r = \begin{pmatrix} 0 & I_r \\ I_r & 0 \end{pmatrix}, \tag{45}$$

with I_r the r-rowed identity matrix. Write

$$B = \begin{pmatrix} B_1 & B_2 \\ 0 & B_3 \end{pmatrix},$$

where B_1 has $2r$ rows and columns. Since B has elements below the diagonal all zero this is true of B_1 and B_3. But

$$B + B' = \begin{pmatrix} B_1 + B_1' & B_2 \\ B_2' & B_3 + B_3' \end{pmatrix} = \begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix},$$

so that $B_2=0$, $B_1+B_1'=W_r$. Then

$$B = \begin{pmatrix} G_1 + \Gamma_r & 0 \\ 0 & G_2 \end{pmatrix}, \qquad \Gamma_r = \begin{pmatrix} 0 & I_r \\ 0 & 0 \end{pmatrix}, \tag{46}$$

and G_1 and $G_2=B_3$ are diagonal matrices. It is clear that r is an invariant of f, and we have proved the following theorem:

THEOREM 16. *Every non-diagonal quadratic form is equivalent to a form*

$$f = \sum_{i=1}^{n} a_i x_i^2 + (x_1 x_{r+1} + \cdots + x_r x_{2r}). \tag{47}$$

Moreover two forms f and g are equivalent only if they have the same rank invariant r.

Let f of the form (47) go into g of the form (47) under a transformation of matrix D. Then D leaves $A+A'$ invariant. We suppose that

$$A = \begin{pmatrix} G_1 + \Gamma_r & 0 \\ 0 & G_2 \end{pmatrix}, \qquad A + A' = \begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix}, \tag{48}$$

and by Lemma 6 see that

$$D(A + A')D' = D\begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix} D' = \begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix} \tag{49}$$

if and only if R is a non-singular matrix, where

$$D = \begin{pmatrix} H & K \\ 0 & R \end{pmatrix}, \qquad HW_rH' = W_r. \tag{50}$$

Then

$$B + N = DAD' = \begin{pmatrix} H(G_1 + \Gamma_r)H' + KG_2K' & KG_2R' \\ RG_2K' & RG_2R' \end{pmatrix} \tag{51}$$

with N alternate, and

$$B = \begin{pmatrix} G_{10} + \Gamma_r & 0 \\ 0 & G_{20} \end{pmatrix}, \tag{52}$$

where G_{20} is a non-alternate symmetric matrix whose diagonal elements coincide with those of RG_2R'. This gives the theorem:

THEOREM 17. *Let*

$$f = \left(\sum_{i=2}^{2r} a_i x_i^2 + x_1 x_{r+1} + \cdots + x_r x_{2r} \right) + \sum_{i=2r+1}^{n} a_i x_i^2$$

and

$$g = \left(\sum_{i=1}^{2r} b_i y_i^2 + y_1 y_{r+1} + \cdots + y_r y_{2r} \right) + \sum_{i=2r+1}^{n} b_i y_i^2 .$$

Then f and g are equivalent in $\mathfrak{F}^{(2)}$ only if the forms

$$\sum_{i=2r+1}^{n} a_i x_i^2, \qquad \sum_{i=2r+1}^{n} b_i y_i^2 \tag{53}$$

are equivalent.

We next write

$$H = \begin{pmatrix} L & U \\ V & M \end{pmatrix}, \tag{54}$$

where L, M, U, V are r-rowed square matrices. Then

$$\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} = H\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}H' = \begin{pmatrix} U & L \\ M & V \end{pmatrix}\begin{pmatrix} L' & V' \\ U' & M' \end{pmatrix} = \begin{pmatrix} UL' + LU' & UV' + LM' \\ VU' + ML' & MV' + VM' \end{pmatrix}, \tag{55}$$

if and only if

$$UL' = (UL')', \qquad MV' = (MV')', \qquad UV' + LM' = I, \tag{56}$$

where I is the r-rowed identity matrix. Then if

$$G_1 = \begin{pmatrix} C & 0 \\ 0 & J \end{pmatrix}, \tag{57}$$

where C and J are r-rowed diagonal matrices, we have

$$\begin{aligned} HG_1H' &= \begin{pmatrix} L & U \\ V & M \end{pmatrix}\begin{pmatrix} C & 0 \\ 0 & J \end{pmatrix}H' = \begin{pmatrix} LC & UJ \\ VC & MJ \end{pmatrix}\begin{pmatrix} L' & V' \\ U' & M' \end{pmatrix} \\ &= \begin{pmatrix} LCL' + UJU' & 0 \\ 0 & VCV' + MJM' \end{pmatrix} + N_0, \end{aligned} \tag{58}$$

where N_0 is alternate. Also

$$\begin{aligned} H\Gamma_r H' &= \begin{pmatrix} L & U \\ V & M \end{pmatrix}\begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix}H' = \begin{pmatrix} 0 & L \\ 0 & V \end{pmatrix}\begin{pmatrix} L' & V' \\ U' & M' \end{pmatrix} = \begin{pmatrix} LU' & LM' \\ VU' & VM' \end{pmatrix} \\ &= \begin{pmatrix} LU' & 0 \\ 0 & VM' \end{pmatrix} + \Gamma_r + \begin{pmatrix} 0 & UV' \\ VU' & 0 \end{pmatrix}, \end{aligned} \tag{59}$$

since $I + LM' = UV'$. The last matrix in (59) is alternate; hence the diagonal matrix G_{10} has the same diagonal elements as

$$KG_2K' + \begin{pmatrix} LCL' + UJU' + LU' & 0 \\ 0 & VCV' + MJM' + VM' \end{pmatrix}. \tag{60}$$

The quadratic form with matrix (60) is a quadratic form $\sum_{i=1}^{2r} b_i y_i^2$. The part corresponding to the arbitrary matrix K is clearly

$$\sum_{j=1}^{2r}\left(\sum_{i=2r+1}^{n} a_i k_{ij}^2\right) y_j^2, \tag{61}$$

for arbitrary k_{ij} and a_i given as in Theorem 17. We now let f and g be two arbitrary forms satisfying the necessary conditions of Theorems 16, 17 so that we may write

$$f = \sum_{i=2r+1}^{n} a_i x_i^2 + \sum_{i=1}^{r} (a_i x_i^2 + x_i x_{i+r} + a_{i+r} x_{i+r}^2), \tag{62}$$

$$g = \sum_{i=2r+1}^{n} a_i y_i^2 + \sum_{i=1}^{r} (b_i y_i^2 + y_i y_{i+r} + b_{i+r} y_{i+r}^2). \tag{63}$$

Also write

$$z = (y_1, \cdots, y_r), \qquad w = (y_{r+1}, \cdots, y_{2r}),$$
$$C = \begin{bmatrix} a_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & a_r \end{bmatrix}, \qquad J = \begin{bmatrix} a_{r+1} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & a_{2r} \end{bmatrix}. \tag{64}$$

Then we have proved the following theorem:

THEOREM 18. *The forms f and g are equivalent in $\mathfrak{F}^{(2)}$ if and only if there exist r-rowed square matrices L, M, U, V such that LU' and VM' are symmetric, $UV'+LM'=I_r$, and the quadratic form*

$$\sum_{i=1}^{2r} b_i y_i^2 - z(LCL' + UJU' + LU')z' - w(VCV' + MJM' + VM')w' \tag{65}$$

may be expressed as (61) *for k_{ij} in $\mathfrak{F}$.*

It does not seem possible to materially simplify (65) for an arbitrary field $\mathfrak{F}^{(2)}$. However we may obtain an analogue of the classical complex number case by proving the theorem:

THEOREM 19. *Every non-diagonal quadratic form with elements in an algebraically closed field $\mathfrak{F}^{(2)}$ of characteristic two and rank invariant r is equivalent to one and only one of the forms*

$$x_1 x_{r+1} + \cdots + x_r x_{2r}, \tag{66}$$

$$x_1 x_{r+1} + \cdots + x_r x_{2r} + x_{2r+1}^2. \tag{67}$$

Hence two non-diagonal quadratic forms are equivalent in $\mathfrak{F}^{(2)}$ if and only if they have the same rank invariant r and the same type (66) *or* (67).

For if $\mathfrak{F}^{(2)}$ is algebraically closed, we use Theorem 15 and transform the form (53) into x_{2r+1}^2 if it is not identically zero. By Theorem 16 our above theorem is true if we can show that the form

$$\sum_{i=1}^{r} (a_i x_i^2 + x_i x_{i+r} + a_{i+r} x_{i+r}^2)$$

is equivalent to (66). This is clearly true if it can be proved that the two forms

$$ax^2 + xy + by^2, \qquad XY$$

are equivalent for every a and b of $\mathfrak{F}^{(2)}$. If $a=b=0$ the result is trivial. We may therefore assume $a\neq 0$ without loss of generality. The equation

$$\omega^2 + \omega + ab = 0$$

has two distinct roots $\lambda, \lambda+1$ in $\mathfrak{F}$. Thus

$$(\omega - \lambda)(\omega - \lambda - 1) \equiv \omega^2 + \omega + ab.$$

Put $\omega = axy^{-1}$ and multiply by $a^{-1}y^2$ to obtain

$$a^{-1}y^2(a^2x^2y^{-2} + axy^{-1} + ab) \equiv ax^2 + xy + by^2 \equiv XY,$$

where

$$X = a^{-1}y(\omega - \lambda) = a^{-1}y(axy^{-1} - \lambda) = x - a^{-1}\lambda y$$

and

$$Y = y(\omega - \lambda - 1) = y(axy^{-1} - \lambda - 1) = ax - y(\lambda + 1).$$

The determinant of the transformation is

$$\begin{vmatrix} 1 & a \\ a^{-1}\lambda & \lambda + 1 \end{vmatrix} = 1,$$

and we have proved our theorem.

III. Pairs of symmetric matrices

1. **The problem.** The theory of the congruence of pairs of symmetric matrices has been studied only in the classical case of matrices with complex elements. The results hold however for matrices with elements in any algebraically closed field whose characteristic is not two. It will be our purpose in the present chapter to develop these results and to show precisely where the classical proofs fail.

2. **The *n*th roots of a matrix.** Let $\mathfrak{J}$ be an integral domain with the property that every two elements of $\mathfrak{J}$ have a greatest common divisor in $\mathfrak{J}$ which

is linearly expressible in terms of them. Then the congruence

$$ax \equiv b \quad (c) \tag{68}$$

has a solution x for every a prime to c.

We consider a prime (or irreducible) element π of $\mathfrak{J}$ and study the congruence

$$f(x) \equiv 0 \quad (\pi^e). \tag{69}$$

Here $f(x)$ is a polynomial in an indeterminate x with coefficients in $\mathfrak{J}$, and e is any positive integer. It is clear that (69) implies that, in particular,

$$f(x) \equiv 0 \quad (\pi) \tag{70}$$

must have a solution. Let then $f(x) \equiv 0$ (π^{e-1}) have a solution x_0, and write $x = x_0 + y\pi^{e-1}$. The Taylor expansion of $f(x)$ then implies that (69) is satisfied if and only if $f(x_0) + yf'(x_0)\pi^{e-1} \equiv 0$ (π^e), that is,

$$yf'(x_0) + q \equiv 0 \quad (\pi), \tag{71}$$

where $\pi^{e-1}q = f(x_0)$. It is clear that (71) has a solution if $f'(x_0) \not\equiv 0$ (π).* This gives the lemma:

LEMMA 8. *Let $f(x) \equiv 0$ (π) have a solution x_0 such that $f'(x_0) \not\equiv 0$ (π). Then there exists a solution of* (69) *for every e.*

The result of our lemma may now be applied to non-singular matrices A with elements in a field $\mathfrak{F}$. Suppose that

$$g = g(\xi) = [\pi_1(\xi)]^{e_1} \cdots [\pi_t(\xi)]^{e_t}$$

is a factorization of the minimum function of A into powers of distinct irreducible functions $\pi_i(\xi)$, and let $\mathfrak{J}$ be the integral domain of all polynomials in the indeterminate ξ. We consider the congruence $x^n \equiv \xi$ (g). This congruence may be easily shown to be solvable modulo g if and only if it is solvable modulo $\pi_i^{e_i}$ for $i = 1, \cdots, t$. Now the derivative of $x^n - \xi$ is $nx^{n-1} \equiv 0$ (π_i) if and only if $nx \equiv 0$ (π_i). But $x^n - \xi \equiv 0$ (π_i) so that either $\xi \equiv 0$ (π_i) or $n \equiv 0$ (π_i). Then $\xi \equiv 0$ (π_i) means that ξ is a factor of $g(\xi)$ which is impossible when A is non-singular. Also $n \not\equiv 0$ (π_i) for π_i irreducible unless the characteristic of $\mathfrak{F}$ divides n. Hence the conditions of our lemma reduce to $x^n - \xi \equiv 0$ (π_i). The equation $\pi_i(\xi) = 0$ defines a field $\mathfrak{F}(\xi_i)$ over $\mathfrak{F}$ equivalent to the set of all residue classes modulo π_i. Then $x^n - \xi \equiv 0$ (π_i) if and only if $\xi_i = x_i^n$ for x_i in

* This is the standard technique for the study of congruences $f(x) \equiv 0$ (mod p) in the theory of numbers (as in L. E. Dickson's *Introduction to the Theory of Numbers*, p. 16, ex. 4). We are using the analogous property of the polynomial domain (cf. Lemma 35.21, p. 60, of MacDuffee's tract on *The Theory of Matrices*).

$\mathfrak{F}(\xi_i)$, and when this condition is satisfied we have $[x(\xi)]^n - \xi \equiv 0 \ (g(\xi))$, $[x(A)]^n = A$. We have proved the following theorem:

THEOREM 20. *Let n be an integer not divisible by the characteristic of $\mathfrak{F}$, A be a non-singular matrix whose minimum function $g(\xi)$ has the distinct irreducible factors $f_i(\xi)$, $\mathfrak{F}_i = \mathfrak{F}(\xi_i)$ be the corresponding algebraic fields over $\mathfrak{F}$. Then there exists a polynomial $P(A)$ whose nth power is A if and only if the equations*

$$\xi_i = x^n \tag{72}$$

have solutions x_i in $\mathfrak{F}(\xi_i)$.

We next consider singular matrices A. Then $g(\xi) = \xi^r g_0(\xi)$, where $g_0(\xi) \not\equiv 0$ (ξ) and $r \geqq 1$. Thus $x^n \equiv \xi \ (g)$ implies that $x^n - \xi \equiv 0 \ (\xi^r)$, $x = \xi x_1$, $\xi^{n-1} x_1^n \equiv 1$ (ξ^{r-1}) which is impossible for $r > 1$. This gives the theorem:

THEOREM 21. *Let $n > 1$, A be a singular matrix whose minimum function $g(\xi)$ is divisible by ξ^2. Then there exists no polynomial in A whose nth power is A.*

As an immediate corollary of the above argument we have the following theorem:

THEOREM 22. *Let $n > 1$, A be a singular matrix whose minimum function has the form $\xi g(\xi)$ where $g(\xi)$ is not divisible by ξ and has irreducible factors with the properties of Theorem 20. Then there exists a polynomial $P(A)$ with coefficients in $\mathfrak{F}$ whose nth power is A.*

Theorem 20 may be applied to prove the following theorem:

THEOREM 23. *Let $\mathfrak{F}$ be an algebraically closed field, n be an integer not divisible by the characteristic of $\mathfrak{F}$, A be a non-singular matrix. Then there exists a polynomial in A with coefficients in $\mathfrak{F}$ whose nth power is A.*

For the fields $\mathfrak{F}_i$ of Theorem 20 are all equal to $\mathfrak{F}$. Moreover $x^n = \xi_i$ has a root x_i in $\mathfrak{F}$; and we have our theorem.

Theorem 23 does not hold for arbitrary matrices A if the characteristic of $\mathfrak{F}$ divides n. For let $\mathfrak{F}$ have characteristic p and B be the p-rowed square matrix all of whose elements are unity. A trivial computation shows that $B^2 = 0$, $B^p = 0$. The matrix $A = I + B$ is non-singular since $A^p = I + B^p = I$. Now any polynomial in A has the form $a_0 + a_1 B$, a_0 and a_1 in $\mathfrak{F}$. Then $(a_0 + a_1 B)^p = a_0^p \neq A$ for any a_0, a_1. If $n = pq$ we have $(a_0 + a_1 B)^n = a_0^n \neq A$. This proves the following theorem:

THEOREM 24. *Let the characteristic of $\mathfrak{F}$ divide n. Then there exist non-singular matrices A such that no polynomial $\phi(A)$ has the property $[\phi(A)]^n = A$.*

3. **Equivalence of pairs of symmetric and alternate matrices.** The result of Theorem 23 may be applied to the theory of equivalence of pairs of sym-

metric matrices. We assume that $\mathfrak{F}$ is algebraically closed as is usual in the classical case. Suppose now that P and Q are non-singular matrices such that

$$PAQ = B, \tag{73}$$

where A and B are either both symmetric or both alternate. Then $(PAQ)' = Q'A'P' = B'$ so that

$$PAQ = Q'AP', \qquad AG' = GA, \tag{74}$$

where

$$G = P^{-1}Q' \tag{75}$$

is non-singular. We now assume that the characteristic of $\mathfrak{F}$ is not two and use Theorem 23 to obtain a polynomial $\phi(G)$ such that

$$[\phi(G)]^2 = G.$$

Now $AG' = GA$ implies that $AG'^2 = G^2A$, $A(G')^k = G^kA$, $A\phi(G') = \phi(G)A$ for any polynomial in G with coefficients in $\mathfrak{F}$. We use the $\phi(G)$ above and have

$$A = \phi(G)A\phi(G')^{-1}, \qquad G^{-1}A = \phi(G)^{-1}A[\phi(G)^{-1}]' = Q'^{-1}PA.$$

Write $H = Q'[\phi(G)]^{-1}$ and obtain

$$HAH' = Q'[\phi(G)]^{-1}A[\phi(G)^{-1}]'Q = Q'Q'^{-1}PAQ = PAQ.$$

We have proved the following theorem:

THEOREM 25. *Let $\epsilon = \pm 1$, $A' = \epsilon A$, P and Q be non-singular matrices such that $(PAQ)' = \epsilon(PAQ)$. Define*

$$G = P^{-1}Q', \qquad H = Q'[\phi(G)]^{-1}, \tag{76}$$

where $\phi(G)$ is a polynomial in G determined so that its square is G. Then $HAH' = PAQ$.

It is clear that the proof we have given of Theorem 25 is not valid for fields of characteristic two. The result itself does not hold. We shall prove this in an example given later.

DEFINITION. *Let A, B, C, D be n-rowed square matrices with elements in $\mathfrak{F}$. Then we say that the pairs (A, B) and (C, D) are equivalent in $\mathfrak{F}$ if there exist non-singular matrices P and Q with elements in $\mathfrak{F}$ such that*

$$PAQ = C, \qquad PBQ = D.$$

We also call (A, B) and (C, D) congruent in $\mathfrak{F}$ if there exists a non-singular matrix H with elements in $\mathfrak{F}$ such that

$$HAH' = C, \qquad HBH' = D.$$

The notion of congruence of pairs may be applied to either symmetric or alternate matrices. It is clear that if A is symmetric, C must be symmetric, and when A is alternate C must be alternate. We have similar necessary conditions on B and D. Then Theorem 25 implies the following theorem:

THEOREM 26. *Let $\mathfrak{F}$ be an algebraically closed field of characteristic not two. Then two pairs of alternate or symmetric matrices satisfying the trivial necessary conditions above are congruent if and only if they are equivalent.*

Conditions that two pairs of matrices be equivalent are expressed in the literature in terms of the invariant factors of the matrices $Ax+B$, $Cx+B$, where x is an indeterminate over $\mathfrak{F}$. This theory holds for an arbitrary $\mathfrak{F}$, and we shall not discuss these known results.

4. **Elementary applications.** There are two simple consequences of the theory of pairs of matrices which seem interesting and appear never to have been noted in the literature. We first let A, B be two non-singular matrices with elements in an algebraically closed field $\mathfrak{F}$ of characteristic not two, and ask the question as to when they are congruent. The answer is given as the corollary:

COROLLARY I. *Write $A=A_1+A_2$, $B=B_1+B_2$ where $A_1=\frac{1}{2}(A+A')\neq 0$ and $B_1=\frac{1}{2}(B+B')\neq 0$ are symmetric, while $A_2=\frac{1}{2}(A-A')\neq 0$ and $B_2=\frac{1}{2}(B-B')\neq 0$ are alternate matrices. Then A and B are congruent if and only if*

$$Ax + A_1, \qquad Bx + B_1$$

have the same invariant factors.

For the theory of invariant factors states that (A, A_1) and (B, B_1) are equivalent if and only if $Ax+A_1$ and $Bx+B_1$ have the same invariant factors. Then $HAH'=B$ implies that $HA'H'=B'$, $H(A\pm A')H'=B\pm B'$, $HA_1H'=B_1$ so that (A, A_1) and (B, B_1) are equivalent when A and B are congruent. Conversely let (A, A_1) and (B, B_1) be equivalent so that

$$PAQ = B, \qquad PA_1Q = B_1.$$

Then $P(A-A_1)Q=B-B_1=B_2=PA_2Q$ and the pairs (A_1, A_2) and (B_1, B_2) are equivalent. By Theorem 26 they are congruent, $HA_1H'=B_1$, $HA_2H'=B_2$, so that $HAH'=H(A_1+A_2)H'=B_1+B_2=B$ as desired.

We next restrict our attention to the field $\mathfrak{C}$ of complex numbers. Corollary I then states that two Hermitian non-singular matrices

$$A = A_1 + A_2 i, \qquad B = B_1 + B_2 i \tag{77}$$

for real A_1, A_2, B_1, B_2 are congruent in $\mathfrak{C}$ if and only if the matrices $Ax+A_1$, $Bx+B_1$ have the same invariant factors. An analogous question arises when

we consider two symmetric matrices (77). Then A_1, A_2, B_1, B_2 are all real symmetric matrices. We then ask for necessary and sufficient conditions that $PA\bar{P}'=B$ for a non-singular P with complex elements. It is clear that then $P\bar{A}'\bar{P}'=\bar{B}'$, $P(A\pm\bar{A}')\bar{P}'=B\pm\bar{B}'$, so that

$$PA_1\bar{P}' = B_1, \qquad PA_2\bar{P}' = B_2.$$

The classical analogue of Theorem 26 states that two pairs of Hermitian complex matrices (A_1, A_2) and (B_1, B_2) are conjunctive if and only if they are equivalent. They are clearly equivalent if and only if (A, A_1) and (B, B_1) are equivalent and we have the statement:

Corollary II. *Two non-singular complex symmetric matrices $A=A_1+A_2i$, $B=B_1+B_2i$ for real symmetric A_1, A_2, B_1, B_2 are conjunctive if and only if*

$$Ax + A_1, \qquad Bx + B_1$$

have the same invariant factors.

5. **Orthogonal equivalence.** The theory of the orthogonal equivalence of two symmetric matrices in an algebraically closed field $\mathfrak{F}$ of characteristic not two is equivalent to the theory of the congruence of pairs (A, B), (C, D) of which A and C are non-singular. The concept of orthogonal equivalence is defined as follows. Let D be a square matrix with elements in a field $\mathfrak{F}$. Then we call D orthogonal if $DD'=I$ is the identity matrix. Clearly then $DD'=D'D$. We consider two symmetric or alternate matrices A, B and call them orthogonally equivalent if $DAD'=B$ for an orthgonal D. Since $B=DAD^{-1}$ is similar to A when A and B are orthogonally equivalent, the condition of similarity is a necessary condition. However we may actually prove the following theorem:

Theorem 27. *Let A and B be both symmetric or both alternate matrices with elements in an algebraically closed field $\mathfrak{F}$ whose characteristic is not two. Then A and B are orthogonally equivalent if and only if they are similar.*

For let $PAP^{-1}=B$. Then $PIP^{-1}=I$ and the pairs (I, A) and (I, B) are equivalent. By Theorem 26 they are congruent, $DID'=I$, $DAD'=B$, D is orthogonal. The converse is trivial.

Theorem 27 is an immediate consequence of Theorem 26. However the connection between the two results is even closer than is indicated by this fact. For assume that the result of Theorem 27 has been proved true independently of Theorem 26. Let now (A, B), (C, G) be any two pairs of matrices such that

$$A' = A, \qquad C' = C, \qquad B' = \epsilon B, \qquad G' = \epsilon G,$$

where $\epsilon = \pm 1$. Assume also that A and C are non-singular. Then (A, B) and (C, G) are congruent if and only if $Ax+B$ and $Cx+G$ have the same invariant factors. This statement of Theorem 26 follows from Theorem 27. We use the existence of a matrix Q such that $QCQ'=I$. Then the invariant factors of $Ax+B$ and $xI+PBP'$ are the same. Similarly those of $Cx+G$ and $xI+QGQ'$ are the same. But when $Ax+B$ and $Cx+G$ have the same invariant factors, so do $xI+QGQ'$, $xI+PBP'$, QGQ' is orthogonally equivalent to PBP'. We let

$$QGQ' = DPBP'D', \qquad DD' = I$$

and have

$$HAH' = Q^{-1}DPAP'D'Q'^{-1} = Q^{-1}Q'^{-1},$$
$$H = Q^{-1}DP, \qquad G = HBH', \qquad C = Q^{-1}Q'^{-1} = HAH'.$$

The above proof cannot, of course, be carried out for fields $\mathfrak{F}$ which are not algebraically closed. For it depends essentially upon the property that $QAQ'=I$ for every non-singular A with elements in $\mathfrak{F}$. However it is clear that in an arbitrary $\mathfrak{F}$ a criterion for the congruence of two pairs always gives a criterion for orthogonal equivalence. For we may take one matrix in each pair to be the identity matrix.

6. ***J*-orthogonal equivalence** The set $\mathfrak{M}_n$ of all n-rowed square matrices with elements in a field $\mathfrak{F}$ is said to possess an involution J over $\mathfrak{F}$ if there is a one-to-one correspondence

$$J: \qquad A \longleftrightarrow A^J \qquad (A, A^J \text{ in } \mathfrak{M}_n)$$

such that

$$(A^J)^J = A, \quad (A+B)^J = A^J + B^J, \quad (AB)^J = B^JA^J, \quad (aI_n)^J = aI_n \tag{78}$$

for every A and B of $\mathfrak{M}_n$ and a of $\mathfrak{F}$. Here I_n is the n-rowed identity matrix. I have proved* that every J is determined† by

$$A^J = E^{-1}A'E, \tag{79}$$

where E is a non-singular matrix $E=\pm E'$. Call A J-symmetric if $A=A^J$, J-alternate if $A=-A^J$ and $\mathfrak{F}$ does not have characteristic two.

Let S be an automorphism of $\mathfrak{M}_n$ over $\mathfrak{F}$. Then there exists a non-singular matrix P such that

$$A^S = P^{-1}AP \tag{80}$$

for every A. The involution

$$A \longleftrightarrow A^{S^{-1}JS} \tag{81}$$

---

* See the author's paper, *Involutorial simple algebras and real Riemann matrices*, loc. cit.

† Note that conversely J determines E only up to a scalar factor.

has been called an involution *cogredient* with J.* In fact

$$A^{S^{-1}} = PAP^{-1}, \qquad (A^{S^{-1}})' = P'^{-1}A'P', \qquad (A^{S^{-1}})^J = E^{-1}P'^{-1}A'P'E,$$

$$A^{S^{-1}JS} = E_0^{-1}A'E_0, \tag{82}$$

where $E_0 = P'EP$ is congruent to E. The set $\mathfrak{M}_n$ may be thought of as the set of linear transformations of a vector space $\mathfrak{R}$, and the replacement of any basis of $\mathfrak{R}$ by any other replaces the matrices of $\mathfrak{M}_n$ by the similar matrices $P^{-1}AP$. They then replace E by $P'EP$, J by $S^{-1}JS$. Hence cogredient involutions are merely different representations of the same abstract involution.

A matrix D is said to be J-orthogonal if

$$D^JD = I_n. \tag{83}$$

Two matrices A and B are called J-orthogonally equivalent if

$$B = D^JAD \tag{84}$$

for a J-orthogonal matrix D. The case where A^J is the transpose of A, that is $E = I_n$, has already been considered in Theorem 27. But we have the following generalization:

THEOREM 28. *Let $\mathfrak{F}$ be an algebraically closed field of characteristic not two, J be an involution of $\mathfrak{M}_n$ over $\mathfrak{F}$, A and B be matrices with elements in $\mathfrak{F}$ such that $A^J = \epsilon A$, $B^J = \epsilon B$, $\epsilon = \pm 1$. Then A and B are J-orthogonally equivalent if and only if they are similar.*

For if A and B are J-orthogonally equivalent they are clearly similar. Conversely let $PAP^{-1} = B$. Now $E^{-1}B'E = \epsilon B$, $E^{-1}A'E = \epsilon A$ so that

$$B_0 = EB, \qquad A_0 = EA, \qquad B_0' = \delta B_0, \qquad A_0' = \delta A_0,$$

where $\delta = \pm 1$ is the product of ϵ and $E'E^{-1} = \pm 1$. Then the pairs (E, A_0) and (E, B_0) have the property $(EPE^{-1})A_0(P^{-1}) = B_0$, $(EPE^{-1})E(P^{-1}) = E$. By Theorem 26 there exists a non-singular matrix D such that

$$D'ED = E, \qquad D'A_0D = B_0.$$

But then

$$D^JD = (E^{-1}D'E)D = I_n,$$

and D is J-orthogonal. Also $B = E^{-1}D'EAD = D^JAD$ is J-orthogonal to A.

---

* N. Jacobson, *A class of normal simple Lie algebras of characteristic zero*, Annals of Mathematics, vol. 38 (1937), pp. 508–517. Note that conversely the property in the preceding footnote implies that E_0 defines an involution cogredient with that defined by E if and only if E_0 is congruent to a scalar multiple of E.

The proof above indicates that the theory of the congruence of pairs of matrices (E, A), (C, B) over any field $\mathfrak{F}$ is equivalent to the theory of J-orthogonal equivalence. We are of course assuming that E and C are non-singular,

$$E' = \epsilon E, \qquad C' = \epsilon C, \qquad A' = \delta A, \qquad B' = \delta B \qquad (\epsilon = \pm 1, \delta = \pm 1).$$

Necessarily C must be congruent to E, and there is no loss of generality if we replace (C, B) by (E, B_1), where B_1 is clearly congruent to B. Then (E, A) and (E, B_1) are congruent if and only if $E^{-1}A$ and $E^{-1}B_1$ are J-orthogonally equivalent. In fact we define $G^J = E^{-1}G'E$ for any matrix G and have

$$A_0 = E^{-1}A, \qquad B_0 = E^{-1}B_1.$$

We then obtain

$$A_0^J = E^{-1}A'E'^{-1}E = \epsilon\delta A_0, \qquad B_0^J = \epsilon\delta B_0.$$

The proof of our theorem then shows that the J-orthogonal equivalence of A_0, B_0 and the congruence of (E, A) and (E, B_1) are equivalent concepts. Notice that this is true for arbitrary fields $\mathfrak{F}$, and that we are not assuming the algebraic closure of $\mathfrak{F}$ or even that the characteristic of $\mathfrak{F}$ is not two.

IV. Similarity of square matrices

1. **Reduction to primary components.** If two matrices A and B are J-orthogonally equivalent, they are similar. It follows that the properties of J-orthogonal equivalence depend essentially upon the theory of the similarity of square matrices. The usual modern formulation of this theory is valid for an arbitrary field* but is in a form unsuited for the application we shall wish to make. Thus we shall give a new formulation.

Our first assumptions from the classical theory are the following lemmas:

LEMMA 9. *Let B be obtained from A by a row permutation followed by the same column permutation. Then B and A are similar.*

LEMMA 10. *Let $f(x) = g(x) \cdot h(x)$ be the characteristic function of an n-rowed square matrix A where $g(x)$ is prime to $h(x)$ and has leading coefficient unity. Then A is similar to*

$$\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}, \tag{85}$$

where G, H have respective characteristic functions $g(x)$, $h(x)$.

We are assuming that G, H, A have elements in a field $\mathfrak{F}$ containing the coefficients of $f(x)$, $g(x)$, $h(x)$, and that our similarity is similarity in $\mathfrak{F}$. The

* For an exposition of the validity of this fact and proofs of the results assumed in this chapter see the author's *Modern Higher Algebra*, chap. 4.

importance of the form (85) is principally due to the fact that for any G and H we have

$$\phi(A) = \begin{pmatrix} \phi(G) & 0 \\ 0 & \phi(H) \end{pmatrix}. \tag{86}$$

This means more precisely that if

$$\phi(x) = x^r + a_1 x^{r-1} + \cdots + a_r \qquad (a_i \text{ in } \mathfrak{F}), \tag{87}$$

then G has m rows, H has $n-m$ rows, and

$$\begin{aligned} &\phi(A) = A^r + a_1 A^{r-1} + \cdots + a_r I_n, \qquad \phi(G) = G^r + \cdots + a_r I_m, \\ &\phi(H) = H^r + \cdots + a_r I_{n-m}. \end{aligned} \tag{88}$$

In particular if $\phi(x)$ is the minimum function of G and $a_r \neq 0$, then

$$G_0 = \begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix}, \qquad \phi(G_0) = \begin{pmatrix} 0 & 0 \\ 0 & a_r I_{n-m} \end{pmatrix}. \tag{89}$$

But then $x\phi(x)$ is the minimum function of the matrix G_0 formed by bordering G by $n-m$ rows and columns of zeros.

We shall call two square matrices *relatively prime if their characteristic functions are relatively prime.* We also say that a square matrix is *primary* if its characteristic function is a power of an irreducible polynomial. Then Lemma 10 gives the following lemma:

LEMMA 11. *Every square matrix A is similar in $\mathfrak{F}$ to a direct sum of relatively prime primary components the product of whose characteristic functions is that of A.*

Let $g(x)$ and $h(x)$ be relatively prime. Then $a(x)g(x)+b(x)h(x)=1$ for polynomials a and b. Define $\delta(x)=a(x)g(x)$, $\gamma(x)=1-\delta(x)$. Then $\gamma(G_0)=I_m$, $\delta(G_0)=0$, $\gamma(H_0)=0$, and $\delta(H_0)=I_{n-m}$, for any m-rowed square matrix G_0 such that $g(G_0)=0$, and any $(n-m)$-rowed H_0 such that $h(H_0)=0$. We use this result in the proof of the following lemma:

LEMMA 12. *Let P be non-singular and*

$$A_0 = P\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}P^{-1} = \begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix}, \tag{90}$$

where the characteristic function $g(x)$ of G and G_0 is prime to the characteristic function $h(x)$ of H and H_0. Then

$$P = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}, \qquad P_1 G P_1^{-1} = G_0, \qquad P_2 H P_2^{-1} = H_0. \tag{91}$$

Write

$$P = \begin{pmatrix} P_1 & P_3 \\ P_4 & P_2 \end{pmatrix}.$$

It is clearly sufficient to prove P_3 and P_4 zero. Form $\gamma(A_0)$ and obtain

$$P\begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}P^{-1} = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$$

so that

$$\begin{pmatrix} P_1 & 0 \\ P_4 & 0 \end{pmatrix} = \begin{pmatrix} P_1 & P_3 \\ 0 & 0 \end{pmatrix},$$

and $P_4=0$, $P_3=0$ as desired.

An evident induction now gives the following lemma:

LEMMA 13. *Let*

$$A = \text{diag}\,[G_1, \cdots, G_t], \quad A_0 = \text{diag}\,[G_{01}, \cdots, G_{0t}] = PAP^{-1}, \tag{92}$$

with relatively prime primary components G_i having the same characteristic functions as G_{0i}. Then P is the direct sum of matrices P_i such that $G_{0i}=P_iG_iP_i^{-1}$.

2. **Indecomposable matrices.** A matrix A is called *indecomposable* in $\mathfrak{F}$ if A is not similar in $\mathfrak{F}$ to a direct sum of two matrices. We shall assume the known lemma:*

LEMMA 14. *A matrix A is indecomposable in $\mathfrak{F}$ if and only if its characteristic function is equal to the power*

$$f(x) = x^n - (a_1x^{n-1} + \cdots + a_n) = [d(x)]^e \qquad (a_i \text{ in } \mathfrak{F}) \tag{93}$$

of an irreducible polynomial $d(x)$ and coincides with its minimum function. Every indecomposable matrix is similar in $\mathfrak{F}$ to

$$\begin{bmatrix} 0 & 1 & 0 & \cdots 0 \\ 0 & 0 & 1 & \cdots 0 \\ \cdot & \cdot & \cdot & \cdots\cdot \\ 0 & 0 & 0 & \cdots 1 \\ a_n & a_{n-1} & a_{n-2} & \cdots a_1 \end{bmatrix}. \tag{94}$$

Conversely the minimum function of (94) *is its characteristic function $f(x)$ and* (94) *is indecomposable if and only if $f(x)=d(x)^e$ for an irreducible $d(x)$.*

Lemmas 11, 13 reduce the problem of the reduction of a matrix to a direct

* See the author's *Modern Higher Algebra*, chap. 4.

sum of indecomposables under similarity transformations to the case where A is primary. When A is primary and has the characteristic function $d(x)^f$, $d(x)$ irreducible, the invariant factors of $xI-A$ are the characteristic functions of its indecomposable components. Then A is similar in $\mathfrak{F}$ to

$$\text{diag}\ [B_1, \cdots, B_t], \tag{95}$$

where B_i has $d(x)^{e_i}$ as characteristic function and is indecomposable,

$$e_1 + \cdots + e_t = f, \qquad e_1 \geqq e_2 \geqq \cdots \geqq e_t \geqq 1. \tag{96}$$

We shall call the e_i the *indices* of (95). In particular

$$e_1 = e, \tag{97}$$

where $d(x)^e$ is the minimum function of both (95) and B_1.

We now prove the following lemma:

LEMMA 15. *Let $d(x)=c[h(x)]$ be irreducible in $\mathfrak{F}$, A be an n-rowed indecomposable matrix with $c(x)^e$ as minimum function,*

$$h(x) = x^m + b_1x^{m-1} + \cdots + b_m. \tag{98}$$

Then the matrix

$$B = \begin{pmatrix} 0 & I_n & 0 & \cdots & 0 \\ 0 & 0 & I_n & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & I_n \\ \Gamma_m & \Gamma_{m-1} & \Gamma_{m-2} & \cdots & \Gamma_1 \end{pmatrix}, \qquad \Gamma_m = -b_mI_n + A, \qquad \Gamma_i = -b_iI_n \quad (i = 1, \cdots, m-1), \tag{99}$$

is indecomposable and has $d(x)^e$ as characteristic function.

For it is clear, from the fact that all elements of B are polynomials in A, that

$$h(B) = B^m + b_1B^{m-1} + \cdots + b_mI = \text{diag}\ [A, \cdots, A],$$

I the nm-rowed identity matrix. Then $[c(A)]^e=0$ so that $[d(B)]^e=c[h(B)]^e=0$. It follows that the minimum function of B divides $[d(x)]^e$. But $d(x)$ is irreducible; thus it has the form $d(x)^g$, $g\leqq e$. However $c[h(B)]^g=0$ implies that $[c(A)]^g=0$ whence $g\geqq e$. Then $g=e$. The degree of the minimum function $d(x)^e$ of B is the order nm of B, and B is indecomposable.

We next let B be a decomposable primary matrix so that we may assume that B is a direct sum of matrices B_i which are indecomposable. If $d(x)^f$ is the characteristic function of B, and the irreducible polynomial $d(x)=c[h(x)]$ as in Lemma 15, we may assume that each B_i has the form (99). But the b_j

are the same in each B_i, and an evident permutation of the rows and corresponding columns of B carries B into a similar matrix of the form (99), where

$$A = \operatorname{diag}[A_1, \cdots, A_t]$$

has the same indices as B. We state this result as in the following lemma:

LEMMA 16. *Let $d(x) = c[h(x)]$ be irreducible, $h(x)$ be given by* (98), *and B have characteristic function $[d(x)]^f$. Then B is similar in $\mathfrak{F}$ to a matrix* (99), *where A has the characteristic function $c(x)^f$ and the same indices as B.*

3. **Two canonical forms.** A matrix N is called *nilpotent* of index e if $N^e = 0$, $N^{e-1} \neq 0$. The minimum function of N is clearly x^e. Then N is similar in $\mathfrak{F}$ to

$$\operatorname{diag}[N_1, \cdots, N_t],$$

where N_i is nilpotent of index e_i and may be taken to be the e_i-rowed square matrix

$$\begin{pmatrix} 0 & 1 & 0 \cdots 0 \\ 0 & 0 & 1 \cdots 0 \\ \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot \\ 0 & 0 & 0 \cdots 1 \\ 0 & 0 & 0 \cdots 0 \end{pmatrix}. \tag{100}$$

Notice that $N_i = 0$ if $e_i = 1$. Also $e = e_1$, N has order $f = e_1 + \cdots + e_t$. The matrices N_i are indecomposable, and thus a nilpotent matrix is indecomposable if and only if its index is its order. But the indices of N are clearly the respective indices of its nilpotent indecomposable components N_i in the sense in which we defined index above.

We have seen that the terminology of indices which we defined for arbitrary matrices has precise connotations for nilpotent matrices. These connotations are made precise also generally by the following theorem:

THEOREM 29. *Let the characteristic function of B be $d(x)^f$, where $d(x) = x^m + a_1x^{m-1} + \cdots + a_m$, ($a_i$ in $\mathfrak{F}$), is irreducible. Then B is similar in $\mathfrak{F}$ to*

$$\begin{pmatrix} 0 & I_f & 0 & \cdots 0 \\ 0 & 0 & I_f & \cdots 0 \\ \cdot & \cdot & \cdot & \cdots\cdot \\ 0 & 0 & 0 & \cdots I_f \\ \Delta_m & \Delta_{m-1} & \Delta_{m-2} & \cdots \Delta_1 \end{pmatrix}, \quad \Delta_m = N - a_mI_f, \quad \Delta_i = -a_iI_f, \quad (i = 1, \cdots, m-1), \tag{101}$$

where N is a nilpotent matrix whose indices are the same as those of B.

Our theorem is an immediate application of Lemma 16 to the case where $h(x)$ is irreducible, $c(x)=x$, $d(x)=c[h(x)]=h(x)$. Then the matrix A of Lemma 16 has x^f as characteristic function and is our nilpotent matrix N.

The matrix (101) is a canonical form of a primary matrix. For some purposes other forms may be preferable. One such form is given in the following:

THEOREM 30. *Let $d(x)=c(x^{p^k})$ be irreducible,*

$$B_j = \begin{bmatrix} 0 & I_\nu & 0 & \cdots & 0 \\ 0 & 0 & I_\nu & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & I_\nu \\ B_{j-1} & 0 & 0 & \cdots & 0 \end{bmatrix}, \qquad \nu = p^{j-1}m, \qquad (j = 1, \cdots, k), \tag{102}$$

where $B_0=A$ is an m-rowed square matrix with $[c(x)]^f$ as characteristic function. Then $B=B_k$ has $[d(x)]^f$ as characteristic function and the same indices as A.

For proof we take $h(x)=x^p$ in Lemma 16 and use an induction on k. This result is of particular use in case $d(x)$ is an inseparable irreducible polynomial over $\mathfrak{F}$ of characteristic p and is equal to $c(x^{p^k})$, $c(x)$ separable.

4. **The algebra $\mathfrak{F}[B]$.** The algebra $\mathfrak{F}[B]$ of all polynomials in a matrix B with characteristic function $d(x)^f$, $d(x)$ irreducible, contains certain sub-fields and certain nilpotent matrices. We first prove the following theorem:

THEOREM 31. *Let the minimum function of B have the form $d(x)^e$ with $d(x)$ irreducible. Write*

$$d(x) = c(x^\pi), \tag{103}$$

where $c(x)$ is separable and

$$\pi = p^k, \quad or \quad \pi = 1$$

according as $d(x)$ is inseparable over $\mathfrak{F}$ of characteristic p or is separable. Then there exists a nilpotent matrix N of index e such that

$$B^\pi = N + A, \qquad c(A) = 0. \tag{104}$$

Thus $\mathfrak{K}=\mathfrak{F}[A]$ is a separable field and the polynomial $x^\pi - A$ is irreducible in $\mathfrak{F}$.

Our proof depends upon the known lemma:*

LEMMA 17. *Let $c(x)$ be irreducible and separable. Then for every e there exists a polynomial $g_e(x)$ such that $c[g_e(x)]$ is divisible by $c(x)^e$.*

The matrix $T=B^\pi$ has $c(x)^e$ as minimum function and Lemma 17 implies

* Loc. cit. chap. 10. The result trivially follows from our Lemma 8.

the existence of a polynomial $g(T)=A_0$ such that $c(A_0)=0$. Since $c(x)$ is separable and irreducible the algebra $\mathfrak{K}=\mathfrak{F}[A_0]$ is a field. We now prove the following lemma:

LEMMA 18. *The matrix $c(T)=N_0$ is nilpotent of index e and the algebra $\mathfrak{K}[N_0]=\mathfrak{F}[T]$.*

For the minimum function of N_0 is clearly x^e. It remains to prove that $\mathfrak{K}[N_0]$ which is contained in $\mathfrak{F}[T]$ has the same order over $\mathfrak{F}$ and equals $\mathfrak{F}[T]$. It is sufficient to show that $a_0+a_1N_0+\cdots+a_{e-1}N_0^{e-1}=0$ for a_i in $\mathfrak{K}$ if and only if the $a_i=0$. But $a_0N_0^{e-1}=0$, $N_0^{e-1}\neq 0$, $a_0=0$ in $\mathfrak{K}$. Similarly all the other $a_i=0$.

The quantity T now has the form $T=A+N$, where A is in $\mathfrak{K}$, $N=(a_1+a_2N_0+\cdots+a_{e-1}N_0^{e-2})N_0$ is nilpotent. But $c(T)=c(A)+Nc_1(N)$ $=N_0$, $c(A)$ is nilpotent and in $\mathfrak{K}$. It follows that $c(A)=0$, $N_0=Nc_1(N)$, $N_0^g=0$ where g is the index of N. Thus $g\geqq e$. Similarly $g\leqq e$ and N has index e. We next prove the following lemma:

LEMMA 19. *Let $x^{p^k}-a$ be reducible in a field $\mathfrak{K}$ containing a. Then $a=b^p$, b in $\mathfrak{K}$.*

For $x^{p^k}-a=g(x)h(x)$, where the constant term of $g(x)$ is the tth power of a root ξ of $x^{p^k}-a=0$, $g(x)$ has degree t. Hence $\xi^{rs}=b_0$ in $\mathfrak{F}$, where s is prime to p, r is a power of p. But then $ss_1\equiv 1$ (p^k) and $\xi^r=b_1$ in $\mathfrak{K}$. Since $r<p^k$ we have $p^k=rp^\nu$, $b_1^{p^\nu}=a$, $a=b^p$, b in $\mathfrak{K}$.

We may finally show that $x^\pi-A$ is irreducible. For otherwise $A=D^p$, D in $\mathfrak{K}$, $c(A)=c(D^p)=0$, $c(x^p)$ is inseparable and has a root D in the separable field $\mathfrak{K}$. This is impossible, and our proof of Theorem 31 is complete.

The matrix A has $c(x)$ as minimum function and is similar to

$$\operatorname{diag}[G,\cdots,G], \tag{105}$$

where G has $c(x)$ as both minimum function and characteristic function. It is well known that the only matrices commutative with G are polynomials in G with coefficients in $\mathfrak{F}$. If $AB=BA$ and we write $B=(B_{ij})$, we obtain $B_{ij}G=GB_{ij}$, the B_{ij} are in $\mathfrak{F}[G]$. Thus B may be regarded as a matrix with elements in the field $\mathfrak{K}=\mathfrak{F}[G]$.

The matrix N of Theorem 31 is commutative with A and hence may be regarded as a nilpotent matrix of index e and elements in $\mathfrak{K}$. The matrix B is also a matrix over $\mathfrak{K}$ but now has minimum function

$$(x^{p^k}-G)^e. \tag{106}$$

Then the construction of matrices with minimum function $[c(x^{p^k})]^e$ is completed by Theorems 30, 31, and the above.

5. **Fields of matrices.** Let $c(x)$ define a separable field $\mathfrak{K}=\mathfrak{F}[G]$ where G is a matrix in our canonical form given by (94) for the separable irreducible polynomial $c(x)$ used as the $f(x)$ of (93). Write

$$c(x) = (x - \alpha_1) \cdots (x - \alpha_n)$$

with distinct α_i in an algebraic extension of $\mathfrak{F}$. Then the Vandermonde matrix*

$$V = (\alpha_{ij}), \qquad \alpha_{ij} = \alpha_j^{i-1} \qquad (i, j = 1, \cdots, n) \tag{107}$$

is non-singular,

$$VV' = (s_{i+k-2}) \tag{108}$$

has elements $s_j=\sum_{i=1}^n \alpha_i^j$ in $\mathfrak{F}$ and determinant the discriminant of $f(x)$. A simple computation gives

$$V^{-1}GV = \alpha = \operatorname{diag}\,[\alpha_1, \cdots, \alpha_n]. \tag{109}$$

Let H be any n-rowed square matrix with elements in $\mathfrak{F}$, and define

$$H_0 = (VV')^{-1}H = (h_{ij}) \qquad (h_{ij} \text{ in } \mathfrak{F}).$$

Then

$$V^{-1}HV = V'H_0V = (\Lambda(\alpha_i, \alpha_j)) \qquad (i, j = 1, \cdots, n), \tag{110}$$

where

$$\Lambda(x, y) = \sum_{i,j}^{1,\cdots,n} x^{i-1}h_{ij}y^{j-1}.$$

Clearly $HG=GH$ if and only if $\Lambda(\alpha_i, \alpha_j)=0$ for $i\neq j$, and $H=h(G)$, where $h(x)=\Lambda(x, x)$.

The result that $HG=GH$ if and only if H is a polynomial in G is true for other types of matrices G as well as those above. In particular it is true for indecomposable nilpotent matrices. Let now A be an n-rowed square matrix with irreducible minimum function $d(x)$ of degree m. This is the case $e=1$ of our theory, and we do not assume that $d(x)$ is separable. If $d(x)=x$, then $A=0$, and we discard this case. Otherwise A is similar in $\mathfrak{F}$ to the matrix

$$\operatorname{diag}\,[G, \cdots, G],$$

where G is given by (94) for $d(x)$, A is now a q-rowed matrix with elements in a field $\mathfrak{F}[G]$. The only matrices commutative with A are q-rowed matrices

* The method of proof of this section has been used many times in the theories of Riemann matrices and of linear transformations. For a partial list of references see the bibliography in the paper referred to in the first footnote on page 387.

with elements in $\mathfrak{F}[G]$, and the minimum function of A over $\mathfrak{F}[G]$ is $x-G$. However the analogous result cannot be obtained when $e>1$ since we cannot in general prove the existence of an inseparable sub-field $\mathfrak{K}$ of our algebra of polynomials in a matrix.

6. **Consequences of the canonical forms.** The transpose A' of any matrix A is similar to A. For certain simple matrices the transformation carrying A into A' assumes an interesting and simple form. We let A have the canonical form

$$A = \begin{pmatrix} 0 & 1 & 0 \cdots 0 \\ 0 & 0 & 1 \cdots 0 \\ \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot \\ 0 & 0 & 0 \cdots 1 \\ a & 0 & 0 \cdots 0 \end{pmatrix} \tag{111}$$

so that the characteristic and minimum functions of A are $f(x)=x^n-a$. Write

$$U = \begin{pmatrix} 0 & 0 \cdots 0 & 1 \\ 0 & 0 \cdots 1 & 0 \\ \cdot & \cdot\cdot\cdot\cdot\cdot & \cdot \\ 0 & 1 \cdots 0 & 0 \\ 1 & 0 \cdots 0 & 0 \end{pmatrix}. \tag{112}$$

Then

$$UA = \begin{pmatrix} a & 0 \cdots 0 & 0 \\ 0 & 0 \cdots 0 & 1 \\ 0 & 0 \cdots 1 & 0 \\ \cdot & \cdot\cdot\cdot\cdot\cdot & \cdot \\ 0 & 1 \cdots 0 & 0 \end{pmatrix}, \qquad UA^2 = \begin{pmatrix} 0 & a & 0 \cdots 0 & 0 \\ a & 0 & 0 \cdots 0 & 0 \\ 0 & 0 & 0 \cdots 0 & 1 \\ 0 & 0 & 0 \cdots 1 & 0 \\ \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot & \cdot \\ 0 & 0 & 1 \cdots 0 & 0 \end{pmatrix}. \tag{113}$$

By an evident computation if

$$f(A) = a_0 + a_1A + \cdots + a_{n-1}A^{n-1}, \tag{114}$$

then

$$Uf(A) = \begin{pmatrix} a_1a & a_2a \cdots a_{n-1}a & a_0 \\ a_2a & a_3a \cdots a_0 & a_1 \\ \cdot & \cdot\ \cdot\cdot\cdot\cdot & \cdot \\ a_{n-1}a & a_0 \cdots a_{n-3} & a_{n-2} \\ a_0 & a_1 \cdots a_{n-2} & a_{n-1} \end{pmatrix}. \tag{115}$$

It is clear that UA is symmetric, and since U is symmetric and non-singular,

$$UAU^{-1} = A'. \tag{116}$$

We shall require (115) later. We shall also use the following existence theorem which is a consequence of Theorem 30 for the case $p=2$:

THEOREM 32. *Let $n=2^k$, $f(x)=x^n-a$ be irreducible in $\mathfrak{F}$. Then there exists an n-rowed square matrix A with $f(x)$ as minimum function and such that*

$$E^{-1}A'E = -A, \qquad S^{-1}A'S = A \tag{117}$$

for an alternate matrix E and a symmetric matrix S.

The theorem is true for $n=2$, $k=1$ since

$$E = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & 1 \\ a & 0 \end{pmatrix}$$

satisfy (117) by direct computation. Assume the theorem true for $k-1$, write $m=2^{k-1}$, and have $S_m^{-1}A_m'S_m=A_m$. By Theorem 30 the matrix

$$A = \begin{pmatrix} 0 & I_m \\ A_m & 0 \end{pmatrix} \tag{118}$$

has $f(x)$ as minimum (and characteristic) function. Then the matrices

$$E = \begin{pmatrix} 0 & S_m \\ -S_m & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & S_m \\ S_m & 0 \end{pmatrix} \tag{119}$$

are alternate and symmetric respectively, and if $\epsilon=\pm 1$,

$$\begin{gathered} \begin{pmatrix} 0 & S_m \\ \epsilon S_m & 0 \end{pmatrix}\begin{pmatrix} 0 & I_m \\ A_m & 0 \end{pmatrix} = \begin{pmatrix} S_mA_m & 0 \\ 0 & \epsilon S_m \end{pmatrix}, \\ \begin{pmatrix} 0 & A_m' \\ I_m & 0 \end{pmatrix}\begin{pmatrix} 0 & S_m \\ \epsilon S_m & 0 \end{pmatrix} = \begin{pmatrix} \epsilon S_mA_m & 0 \\ 0 & S_m \end{pmatrix}, \end{gathered} \tag{120}$$

since $A_m'S_m=S_mA_m$. Put $\epsilon=-1$ and obtain $EA=-A'E$. The value $\epsilon=1$ gives $SA=A'S$ as desired.

V. ORTHOGONAL EQUIVALENCE IN $\mathfrak{F}$ OF CHARACTERISTIC TWO

1. **The problem.** Our principal interest will be in obtaining a complete determination of the invariant factors of any symmetric matrix whose elements are in a field of characteristic two.* Part of this theory will be concerned with the orthogonal equivalence of symmetric matrices, and we shall

* The field of reference throughout this chapter will be any field of characteristic two and we shall drop the notation $\mathfrak{F}^{(2)}$ and simply use $\mathfrak{F}$.

show why it can happen that two symmetric matrices may be similar but not orthogonally equivalent.

2. ***J*-orthogonal equivalence.*** Let J be an involution over $\mathfrak{F}$ of the algebra $\mathfrak{M}_n$ of all n-rowed square matrices with elements in a field $\mathfrak{F}$ (with any characteristic). Suppose that

$$A^J = \epsilon A, \qquad B = PAP^{-1} \qquad (\epsilon = \pm 1) \tag{121}$$

for a non-singular matrix P. We may then prove the lemma:

LEMMA 20. *The matrix B of* (121) *has the property*

$$B^J = \epsilon B \tag{122}$$

if and only if P^JP is commutative with A.

For $B^J = (P^J)^{-1}\epsilon AP^J = \epsilon B = \epsilon PAP^{-1}$, $P^JPA = AP^JP$. The converse is trivial.

If B is any matrix similar to A so that $B = PAP^{-1}$, then $B = DAD^{-1}$ if and only if $D = PC$ where C is commutative with A. Then

$$D^JD = C^J(P^JP)C. \tag{123}$$

But D is J-orthogonal if and only if D^JD is the n-rowed identity matrix.

LEMMA 21. *The matrix PAP^{-1} is J-orthogonally equivalent to A if and only if P^JP is congruent to the identity matrix under a transformation* (123) *with C commutative with A.*

Let us study the implications of Lemma 21 in a special case. Assume that A is such that the only matrices commutative with A are polynomials in A with coefficients in $\mathfrak{F}$. Then $P^JP = \phi(A)$ is such a polynomial and $C = C(A)$. If $A^J = A$, then $C^J = C$, and $C^J\phi C = I$ if and only if

$$\phi = C^{-2} \tag{124}$$

is the square of a polynomial in A. Conversely if $\phi = Q^JQ = Q^2$, we have $P^JP = Q^JQ$, $D = PQ^{-1}$ is J-orthogonal, and

$$B = PAP^{-1} = DAD^{-1} \tag{125}$$

is J-orthogonally equivalent to A.

We shall see in an example later that it is possible for a given $\phi(A)$ to have the form $P'P$ but not the form $[Q(A)]^2$. Thus the restriction above is not a redundant consequence of the equation $\phi(A) = P'P$.

* For recent results implying theorems on the J-orthogonal equivalence of matrices over an arbitrary field of characteristic not two, see the papers of John Williamson in the American Journal of Mathematics, vol. 57 (1935), pp. 475–490; vol. 58 (1936), pp. 141–163; and vol. 59 (1937), pp. 399–413.

In the remainder of the chapter we shall assume, unless it is otherwise stated, that $\mathfrak{F}$ has characteristic two. We shall also restrict our attention to the case of ordinary orthogonal equivalence.

3. **Separable fields.** Every separable field $\mathfrak{K}$ over $\mathfrak{F}$ is a simple extension

$$\mathfrak{K} = \mathfrak{F}[A] \tag{126}$$

of all polynomials in A with coefficients in $\mathfrak{F}$, where A is a root of an irreducible separable equation

$$f(x) = x^m - (a_1x^{m-1} + \cdots + a_m) = 0. \tag{127}$$

We may easily prove the following theorem:

THEOREM 33. *Let $f(x)$ of* (127) *be separable and irreducible in $\mathfrak{F}$ of characteristic two. Then there exists an m-rowed symmetric matrix G with elements in $\mathfrak{F}$ and $f(x)$ as characteristic function.*

For let

$$G_0 = \begin{pmatrix} 0 & 1 & 0 & \cdots 0 \\ 0 & 0 & 1 & \cdots 0 \\ \cdot & \cdot & \cdot & \cdots\cdot \\ 0 & 0 & 0 & \cdots 1 \\ a_m & a_{m-1} & a_{m-2} & \cdots a_1 \end{pmatrix}. \tag{128}$$

Then every m-rowed square matrix with $f(x)$ as characteristic function is indecomposable, has $f(x)$ as minimum function, and is similar to G_0. In (109) we saw that $V^{-1}G_0V$ is a diagonal matrix and is thus symmetric. Thus $V'G_0'V'^{-1} = V^{-1}G_0V$,

$$VV'(G_0') = G_0(VV'). \tag{129}$$

The diagonal elements of VV' are

$$s_{i+i-2} = \sum_{j=1}^{m} \alpha_j^{2i-2} = \left(\sum_{j=1}^{m} \alpha_j^{i-1}\right)^2 = (s_{i-1})^2. \tag{130}$$

Hence VV' is a definite matrix. It is this remarkable property which gives us our result, Theorem 33.

We may now write $VV' = RR'$, where R is a non-singular matrix with elements in $\mathfrak{F}$. Then $RR'G_0' = G_0RR'$, $R^{-1}G_0R = R'G_0'R'^{-1}$. Define

$$G = R^{-1}G_0R, \tag{131}$$

and obtain $G' = R'G_0'R'^{-1} = G$. The matrix G is symmetric and is similar to G_0. It is the desired matrix. Notice that

$$W = V^{-1}R, \qquad W'W = I, \qquad WGW^{-1} = \alpha,$$

so that G is orthogonally equivalent in an algebraic extension of $\mathfrak{F}$ to the diagonal matrix α. But then we may prove the lemma:

LEMMA 22. *Let $\mathfrak{F}_0 = \mathfrak{F}[G]$ be the field of symmetric matrices consisting of all polynomials with coefficients in $\mathfrak{F}$ of the matrix G of* (131). *Then the only alternate matrix of $\mathfrak{F}_0$ is the zero matrix. Moreover a matrix of $\mathfrak{F}_0$ is definite if and only if it is the square of a non-zero quantity of $\mathfrak{F}_0$.*

For if $\psi(G)$ is in $\mathfrak{F}_0$, we have

$$W\psi(G)W' = \psi(\alpha).$$

If $\psi(G)$ is alternate so is the diagonal matrix $\psi(\alpha)$. But then $\psi(\alpha) = 0$, $\psi(G) = 0$. Let $\psi(G)$ be definite so that we may write $\psi(G) = H'H$. Then $W = W'^{-1} = (R'V'^{-1})^{-1} = V'R'^{-1}$,

$$\psi(\alpha) = W\psi(G)W' = (TV)'(TV),$$

where $T = HR^{-1}$ has elements t_{ij} in $\mathfrak{F}$. The elements in the first column of TV are

$$t_i(\alpha_1) = \sum_{j=1}^{m} t_{ij}\alpha_1^{\,j-1} \qquad (i = 1, \cdots, m).$$

Thus the element in the first row and column of $\psi(\alpha)$ is

$$\psi(\alpha_1) = \sum_{i=1}^{m} [t_i(\alpha_1)]^2 = \left[\sum_{i=1}^{m} t_i(\alpha_1)\right]^2 = [\tau(\alpha_1)]^2,$$

where $\tau(\alpha_1)$ is in $\mathfrak{F}[\alpha_1]$. Then $\psi(G) = [\tau(G)]^2$, $\tau(G) \neq 0$ in $\mathfrak{F}_0$.

The only matrices commutative with G are quantities of $\mathfrak{F}_0 = \mathfrak{F}[G]$, a field of symmetric matrices. We let

$$A = \operatorname{diag}[G, \cdots, G] \tag{132}$$

have $n = qm$ rows, so that A is a q-rowed matrix with elements in $\mathfrak{F}_0$. Let

$$\mathfrak{M}_q \quad \text{over} \quad \mathfrak{F}_0 \tag{133}$$

be the set of all q-rowed matrices with elements in $\mathfrak{F}_0$. Then $\mathfrak{M}_q$ is the algebra of all n-rowed matrices commutative with A. Every such matrix has the form

$$B = (B_{ij}) \qquad (B_{ij} \text{ in } \mathfrak{F}_0,\ i, j = 1, \cdots, q)$$

and is an n-rowed matrix with elements in $\mathfrak{F}$. We define

$$B^T = (B_{0ij}), \qquad B_{0ij} = B_{ji} \qquad (i, j = 1, \cdots, q)$$

so that B^T is the transpose of B considered as a q-rowed matrix of $\mathfrak{M}_q$ over $\mathfrak{F}_0$.

But B is an n-rowed square matrix with elements in $\mathfrak{F}$, and

$$B' = (B_{1ij}), \qquad B_{1ij} = B'_{ji} \qquad (i, j = 1, \cdots, q).$$

However every B_{ji} is symmetric, and thus we have proved that $B^T = B'$ for every B of $\mathfrak{M}_q$ over $\mathfrak{F}_0$. We now have the following result:

THEOREM 34. *Let S be an n-rowed symmetric matrix commutative with A. Then S is in $\mathfrak{M}_q$ over $\mathfrak{F}_0$ and is symmetric if and only if it is symmetric in $\mathfrak{M}_q$, is alternate if and only if it is alternate in $\mathfrak{M}_q$, and is definite if and only if it is definite in $\mathfrak{M}_q$.*

For $S^T = S'$ and $S = S'$ if and only if $S = S^T$. If $S = (S_{ij})$ and the S_{ij} are in $\mathfrak{F}_0$, then $S = S'$ and $S_{ii} = 0$ implies that S is alternate. Conversely if S is alternate, we have $S = S'$, and the S_{ii} are alternate. By the lemma above the $S_{ii} = 0$; S is alternate in $\mathfrak{M}_q$. If S is definite in $\mathfrak{M}_q$, the S_{ii} are the squares of quantities of $\mathfrak{F}_0$ and are thus zero or definite; S is definite. Conversely let S be a definite n-rowed matrix. Then the principal sub-matrices S_{ii} are either zero or definite and at least one $S_{ii} \neq 0$. By our lemma each $S_{ii} = (\tau_{ii})^2$, τ_{ii} in $\mathfrak{F}_0$ and not all zero, S is definite in $\mathfrak{M}_q$.

Consider a symmetric field of matrices $\mathfrak{K} = \mathfrak{F}[B]$, B a root of $f(x) = 0$ of (127). Then B is similar to A and

$$B = PAP^{-1}, \qquad A = \operatorname{diag}[G, \cdots, G], \tag{134}$$

as in (132). Since A and B are symmetric we may apply Lemma 20 to see that $P'P$ is in $\mathfrak{M}_q$ over $\mathfrak{F}_0$. By Theorem 34 the matrix $P'P$ is definite in $\mathfrak{M}_q$ and $P'P = Q'Q$ where Q is in $\mathfrak{M}_q$, $QA = AQ$. Lemma 21 then states that B is orthogonally equivalent to A, and we have the following theorem:

THEOREM 35. *Any two n-rowed symmetric matrices with the same separable irreducible minimum function over $\mathfrak{F}$ of characteristic two are orthogonally equivalent in $\mathfrak{F}$.*

Theorem 34 may be applied to prove a generalization of Theorem 32.

THEOREM 36. *Let $f(x) = c(x^{2^k})$ be irreducible in $\mathfrak{F}$ of characteristic two, $c(x)$ be separable of degree m, and $q = 2^k > 1$. Then there exists an alternate n-rowed square matrix E and an n-rowed matrix B such that*

$$f(B) = 0, \qquad E^{-1}B'E = B, \tag{135}$$

for every n divisible by $2^k m$.

It is clearly sufficient to give the proof for $n = 2^k m$. We construct an m-rowed symmetric matrix G which is a root of $c(x) = 0$. Use Theorem 32 for the field $\mathfrak{F}_0$, and $a = G$ and obtain an n-rowed square matrix B such that B

considered as a matrix of 2^k rows over $\mathfrak{F}_0$ has $x^{2^k}-G$ as minimum function. Then $B^{2^k}=A$, A the matrix of (132), $f(B)=0$ as desired. Also there exists a matrix E such that $E^{-1}B^JE=B$, $E^J=E$, where E is alternate, J is the involution of the algebra of all 2^k-rowed matrices over $\mathfrak{F}_0$. By Theorem 34 we have E alternate, $E^{-1}B'E=B$.

4. **Application to primary matrices.** Let S be a symmetric primary matrix. We apply Theorem 31 to prove the existence of a polynomial $A=A(S)$ in S such that

$$S^{2k} = A + N, \tag{136}$$

where N is nilpotent, $c(A)=0$. Here the minimum function of S is $[c(x^{2^k})]^e=0$; N has index e. By an orthogonal transformation we may transform A into the form (132). By Theorem 34 both N and S are symmetric over $\mathfrak{F}_0$, N is nilpotent. Then $S^{2^k}=GI_q+N$, and the minimum and characteristic functions of S over $\mathfrak{F}_0$ are

$$(x^{2k} - G)^e, \qquad (x^{2k} - G)^f, \tag{137}$$

where $x^{2^k}-G$ is irreducible in the field $\mathfrak{F}_0$ and S is now a matrix of $q=2^kf$ rows.

Let T be symmetric and similar to S. There is no loss of generality if we take $A(T)$, which is similar to A, equal to A. For by Theorem 35 the matrix $A(T)$ is orthogonally equivalent to A. But then S and T are matrices with elements in $\mathfrak{F}_0$. We now prove the statement:

Theorem 37. *The matrices S and T are orthogonal in $\mathfrak{F}$ if and only if they are orthogonal considered as matrices over $\mathfrak{F}_0$.*

For $SA=AS$, $T=PSP^{-1}$, so that $P'P$ is commutative with S and hence with A. Then $P'P$ is definite, and we have already seen that $P'P$ is a definite matrix over $\mathfrak{F}_0$. Now S and T are orthogonally equivalent if and only if $C'(P'P)C=I$ for $CS=SC$. Thus $CA=AC$, C is a matrix with elements in $\mathfrak{F}_0$ and is commutative with S; S and T are orthogonally equivalent when considered as matrices over $\mathfrak{F}_0$.

We now apply Theorem 30 to see that S is similar in $\mathfrak{F}_0$ to a matrix S_k, where

$$S_j = \begin{pmatrix} 0 & I_{\nu_j} \\ S_{j-1} & 0 \end{pmatrix}, \qquad \nu_j = 2^{j-1}m \qquad (j = 1, \cdots, k), \tag{138}$$

and

$$S_1 = \begin{pmatrix} 0 & I_{\nu_0} \\ G + N_1 & 0 \end{pmatrix}. \tag{139}$$

By Theorem 30 the matrix N_1 is nilpotent and has the same indices as S. Finally Theorem 37 implies that T is orthogonally equivalent to S over $\mathfrak{F}$ if and only if T is orthogonally equivalent to S over $\mathfrak{F}_0$. We have thus reduced our considerations for primary matrices to the case of matrices with minimum function $(x^{2^k}-a)^e$, a in our field $\mathfrak{F}$. We shall prove that $e>1$; that is, the following theorem:

THEOREM 38. *There exist no inseparable fields of symmetric matrices over $\mathfrak{F}$ of characteristic two.*

For let $\mathfrak{K}$ over $\mathfrak{F}$ be inseparable. Then there exists a symmetric matrix S in $\mathfrak{K}$ such that $S^2=A$, A as in (132). Then

$$S^2 = \operatorname{diag}[G, \cdots, G],$$

where x^2-G is irreducible in $\mathfrak{F}$. But S^2 is definite over $\mathfrak{F}$ since $S^2=S'S$, S is non-singular. By the proof of Lemma 22 the matrix S^2 is definite over $\mathfrak{F}_0$, $G=[\phi(G)]^2$, x^2-G is reducible in $\mathfrak{F}$, and we have a contradiction.

5. **Nilpotent matrices.** We shall construct symmetric nilpotent matrices with arbitary indices. Let U be defined as in (112), N_0 be the matrix A of (111) with $a=0$. Then N_0 is nilpotent of index and order n and is indecomposable.

The matrix

$$(140)\qquad Uf(N_0) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & a_0 \\ 0 & 0 & \cdots & 0 & a_0 & a_1 \\ 0 & 0 & \cdots & a_0 & a_1 & a_2 \\ \cdot & \cdot & \cdot\cdot\cdot & \cdot & \cdot & \cdot \\ 0 & a_0 & \cdots & a_{n-4} & a_{n-3} & a_{n-2} \\ a_0 & a_1 & \cdots & a_{n-3} & a_{n-2} & a_{n-1} \end{pmatrix},$$

for any polynomial $f(N)$ with coefficients a_i in $\mathfrak{F}$. We may write

$$f(x) \equiv f_1(x^2) + xf_2(x^2).$$

Then $f_i(x^2)$ is the square of a polynomial in x with coefficients in $\mathfrak{F}$ if and only if its coefficients are squares in $\mathfrak{F}$. We now have the theorem:

THEOREM 39. *Let $h(x)=f_1(x^2)$ or $f_2(x^2)$ according as n is odd or even. Then $Uf(N_0)$ is alternate if and only if $h(x)\equiv 0$ and is semi-definite if and only if $h(x)=[g(x)]^2\not\equiv 0$.*

For we observe that the only elements on the main diagonal of $Uf(N_0)$ are zeros and the complete set of coefficients of $h(x)$. Thus $Uf(N_0)$ alternate implies that $h(x)\equiv 0$; $Uf(N_0)$ semi-definite implies that $h(x)$ is a square.

The matrices $f(N_0)$, $Uf(N_0)$ are non-singular if and only if $a_0 \neq 0$. In particular $U(I+N_0)$ is non-singular. By Theorem 39 it is definite. We now write

$$U(I+N_0) = P'P, \qquad N = PN_0P^{-1}.$$

Since $UN_0 = N_0'U$ we have $P'PN_0 = N_0'P'P$, $PN_0P^{-1} = N = P'^{-1}N_0'P' = N'$, and N' is symmetric.

We have thus proved the existence of an indecomposable nilpotent symmetric matrix N of any order n. Every nilpotent matrix is similar to a direct sum of indecomposable nilpotent matrices, we form such a direct sum and obtain the result:

THEOREM 40. *There exist symmetric nilpotent matrices with any indices.*

6. **Primary matrices.** We shall require the following lemma:

LEMMA 23. *Let $a \neq 0$ be in $\mathfrak{F}$, $e_1 \geqq e_2 \geqq \cdots \geqq e_t \geqq 1$, be integers such that $e_1 > 1$. Then there exists a symmetric nilpotent matrix N with $e_1, \cdots, e_t$ as indices and a symmetric matrix Q commutative with N such that*

$$\begin{pmatrix} Q & 0 \\ 0 & Q(a+N) \end{pmatrix} \tag{141}$$

is semi-definite.

For let us take

$$N = \begin{pmatrix} N_1 & 0 \\ 0 & N_2 \end{pmatrix},$$

where the index of N_1 is $e_1 > 1$. Then we choose $N_1 \neq 0$ to have order e_1 and be indecomposable, N_1 to be the matrix

$$PN_0P^{-1}, \qquad P'P = U(I+N_0)$$

as in §5. If

$$Q = \begin{pmatrix} Q_1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad Q_1 = f(N_1),$$

then Q is symmetric and commutative with N. Moreover (141) is semi-definite if and only if

$$\begin{pmatrix} Q_1 & 0 \\ 0 & Q_1(a+N_1) \end{pmatrix} \tag{142}$$

is semi-definite. The matrix (142) is congruent to

$$\begin{pmatrix} P'Q_1P & 0 \\ 0 & P'Q_1(a+N_1)P \end{pmatrix}. \tag{143}$$

Now $P'f(N_1)P=P'PP^{-1}f(N_1)P=P'Pf(N_0)=U(I+N_0)f(N_0)$. If e_1 is even, we take $Q_1=I+N_1$ and have

$$P'f(N_1)P = U(I + N_0)(I + N_0) = U(I + N_0^2)$$

alternate by Theorem 39. Also

$$P'Q_1(a + N_1)P = U(I + N_0^2)(a + N_0) = U(a + aN_0^2 + N_0 + N_0^3)$$

is semi-definite. But then the matrix (143) congruent to (142) is semi-definite, and so is (141). Similarly when e_1 is odd we take $f(N_1)=Q_1=N_1+N_1^3$,

$$P'Q_1P = U(N_0 + N_0^3), \quad P'Q_1(a + N_1)P = U(N_0^2 + N_0^4 + aN_0 + aN_0^3).$$

The first of these matrices is alternate, the second is semi-definite since now $e_1 \geqq 3$, and $N_0^2 \neq 0$.

We next prove the lemma:

LEMMA 24. *Let M be a symmetric matrix with elements in an infinite field $\mathfrak{F}$, and let Q be symmetric and commutative with M such that*

$$\begin{pmatrix} Q & 0 \\ 0 & QM \end{pmatrix} \tag{144}$$

is semi-definite. Then the matrix

$$B_0 = \begin{pmatrix} 0 & M \\ I & 0 \end{pmatrix} \tag{145}$$

is similar to a symmetric matrix B, and there exists a matrix $S=S'$ commutative with B such that

$$\begin{pmatrix} S & 0 \\ 0 & SB \end{pmatrix} \tag{146}$$

is semi-definite.

Since $\mathfrak{F}$ is infinite we apply Theorem 8 to obtain a quantity $b\neq 0$ in $\mathfrak{F}$ such that the matrix b^4I+Q^2M is definite. Then

$$Q_0 = \begin{pmatrix} Q & b^2 \\ b^2 & QM \end{pmatrix} \tag{147}$$

is semi-definite since (144) is semi-definite. But

$$|Q_0| = \begin{vmatrix} Q & b^2I \\ b^2I & QM \end{vmatrix} = \begin{vmatrix} 0 & b^2I + b^{-2}Q^2M \\ b^2I & QM \end{vmatrix} = |b^4I + Q^2M| \neq 0, \tag{148}$$

so that Q_0 is a definite symmetric matrix, and we may write

$$Q_0 = P_0' P_0, \tag{149}$$

where P_0 is non-singular. Also

$$Q_0B_0 = \begin{pmatrix} b^2 & QM \\ QM & b^2M \end{pmatrix}, \qquad B_0'Q_0 = Q_0B_0 \tag{150}$$

by direct computation. Then $P_0'P_0B_0 = B_0'P_0'P_0$.

The matrix $B = P_0B_0P_0^{-1}$ is now the desired symmetric matrix. Take $S = (P_0^{-1})'S_0P_0^{-1}$ congruent to S_0 given by

$$\begin{pmatrix} Q & Q \\ Q & QM \end{pmatrix}. \tag{151}$$

Then S_0 is semi-definite and so is S. The matrix

$$S_0B_0 = S_0\begin{pmatrix} 0 & M \\ I & 0 \end{pmatrix} = \begin{pmatrix} Q & QM \\ QM & QM \end{pmatrix} = B_0'S_0 \tag{152}$$

is also semi-definite. Now $SB = BS = (P_0')^{-1}S_0P_0^{-1}P_0B_0P_0^{-1} = (P_0^{-1})'(S_0B_0)P_0^{-1}$ is congruent to S_0B_0 and is semi-definite. This proves that (146) is semi-definite and completes the proof of our lemma.

We use Lemma 24 to prove a fundamental result:

THEOREM 41. *A primary matrix B with elements in a field $\mathfrak{F}$ of characteristic two is similar in $\mathfrak{F}$ to a symmetric matrix if and only if the minimum function of B is not an irreducible inseparable polynomial.*

For if B has an irreducible inseparable minimum function and B_0 is symmetric and similar to B, the field $\mathfrak{F}[B_0]$ is inseparable, contrary to Theorem 38. Conversely let the minimum function of B be not an irreducible inseparable polynomial. Then the argument of §4 together with Theorem 40 reduces the proof of our theorem to the question of the existence of a symmetric matrix B with arbitrary indices

$$e = e_1 \geqq e_2 \geqq \cdots \geqq e_t \geqq 1, \qquad e > 1, \tag{153}$$

and with

$$(x^{2^k} - a)^e \qquad (a \text{ in } \mathfrak{F}) \tag{154}$$

as minimum function, $x^{2^k} - a$ irreducible in $\mathfrak{F}$. By Theorem 40 there exists a symmetric nilpotent matrix N with indices $e_1, \cdots, e_t$. The matrix $M = aI_f + N$ has the same indices as N, where $f = e_1 + \cdots + e_t$, and so does

$$\begin{pmatrix} 0 & M \\ I_f & 0 \end{pmatrix} \tag{155}$$

by Theorem 30. We use Lemma 24 and choose the symmetric matrix M so that there exists an M_2 similar to (155) and having the further property of Lemma 24. By Theorem 30 the matrix M_2 has the same indices as N. Its characteristic function is $(x^2-a)^f$. An evident induction yields a matrix M_k which is symmetric and is the desired matrix B.

A field $\mathfrak{F}$ which is perfect has the property that there are no inseparable polynomials over $\mathfrak{F}$. For example every finite field is perfect. But then Lemma 11 and Theorem 41 give the following theorem:

THEOREM 42. *Every square matrix with elements in a perfect field $\mathfrak{F}$ of characteristic two is similar in $\mathfrak{F}$ to a symmetric matrix.*

7. **Orthogonal equivalence of primary matrices.** We have reduced the problem of the orthogonal equivalence of primary symmetric matrices to the case of matrices with characteristic function $(x^{2^k}-a)^f$, $x^{2^k}-a$ irreducible in $\mathfrak{F}$, The case $a=0$ is the case of nilpotent matrices. We may now easily show that two matrices may be similar and not orthogonally equivalent. Since the first part of our problem is that of the orthogonal equivalence of symmetric nilpotent matrices, and since, even in this case, the only criteria are of a necessarily complicated nature, we shall restrict our attention to the nilpotent case. It is the only case occurring when $\mathfrak{F}$ is perfect.

Every symmetric nilpotent matrix A with indices $e_1, \cdots, e_t$ is similar in $\mathfrak{F}$ to

$$N = \operatorname{diag}[N_1, \cdots, N_t], \tag{156}$$

where N_i is nilpotent of index and order e_i and may be taken to be symmetric by Theorem 41. Then N is symmetric and

$$A = LNL^{-1}, \qquad L'L = Q, \tag{157}$$

is commutative with N by Lemma 20. This last condition states that

$$Q = (Q_{ij}), \qquad Q_{ij}N_j = N_iQ_{ij} \qquad (i, j = 1, \cdots, t), \tag{158}$$

so that in particular the Q_{ii} are polynomials in N_i. Also by Lemma 21 if

$$A_0 = L_0NL_0^{-1}, \qquad Q_0 = L_0'L_0, \tag{159}$$

and A_0 is similar to A, then A_0 is orthogonally equivalent to A if and only if $Q_0=C'QC$ is congruent to Q under a transformation whose matrix C is commutative with N. These are the formal orthogonal equivalence conditions.

We first study the case where $N=N_1$ is indecomposable. By the proof of Theorem 40 we may choose N so that

$$P'P = U(I + N_0), \qquad N = PN_0P^{-1}, \tag{160}$$

with (140) holding. Now

$$Q = Q(N), \qquad Q_0 = Q_0(N), \qquad C = C(N) \tag{161}$$

are all polynomials in N and our condition is simply

$$Q_0 = C^2 Q. \tag{162}$$

The matrix Q is definite if and only if $P'QP = P'PQ(N_0) = U(I+N_0)Q(N_0)$ is definite. Write

$$f(N_0) = (I + N_0)Q(N_0). \tag{163}$$

Since $I+N_0$ is non-singular there exists a $Q(N_0)$ for any $f(N_0)$. But Theorem 39 gives necessary and sufficient conditions that $Uf(N_0)$ be definite. Moreover if $f_0(N_0) = (I+N_0)Q_0$, then $Uf_0(N_0)$ may be definite for many polynomials $Q_0 \neq QC^2$. In fact if $f = f_1 + N_0 f_2$, then $fC^2 = f_1C^2 + N_0 f_2 C^2 = f_{10} + N_0 f_{20}$ if and only if $f_{10} = f_1 C^2, f_{20} = f_2 C^2$. For example, if $e = 3$, we may explicitly compute $Uf(N_0)$ which is definite if and only if $a_0 = b_0^2 \neq 0$, $a_2 = b_2^2$,

$$f(N_0) = (b_0 + b_1 N)^2 + a_2 N_0.$$

We write $c(N_0) = c_0 + c_1 N_0 + c_2 N_0^2$, $[c(N_0)]^2 = c_0^2 + c_1^2 N_0^2$, so that

$$[c(N_0)]^2 f(N_0) = [b_0 c_0 + (c_0 b_1 + c_1 b_0) N_0]^2 + c_0^2 a_2 N_0.$$

Now $f_0(N_0) = (d_0 + d_1 N)^2 + d_2 N_0 = [c(N_0)]^2 f(N_0)$ for some $c(N_0)$ if and only if $d_0 = b_0 c_0$, $d_1 = c_0 b_1 + c_1 b_0$, $d_2 = c_0^2 a_2$. Then $c_0 = d_0 b_0^{-1}$ and $c_1 = (d_1 - c_0 b_1) b_0^{-1}$ are determined. But d_2 is at our choice and can be given distinct from $c_0^2 a_2$. Notice that in general our first condition is $f_{10} = c^2 f_1$ which may always have a solution c, but that this solution may not satisfy $f_{20} = f_2 c^2$.

THEOREM 43. *There exist similar indecomposable symmetric nilpotent matrices which are not orthogonally equivalent.*

Decomposable nilpotent matrices need not be orthogonally decomposable. A very simple example may be obtained as follows. Let

$$E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad A = \begin{pmatrix} E & E \\ E & E \end{pmatrix}.$$

The matrix A is a nilpotent matrix of index two in $\mathfrak{F}$ of characteristic two and is similar in $\mathfrak{F}$ to

$$\begin{pmatrix} N & 0 \\ 0 & N \end{pmatrix}, \qquad N = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

If A were orthogonally equivalent to a direct sum, the components would be necessarily nilpotent of index two, hence A would be orthogonal to

$$\begin{pmatrix} aN & 0 \\ 0 & bN \end{pmatrix}, \tag{164}$$

a, b not zero and in $\mathfrak{F}$. For the only two-rowed nilpotent symmetric matrices are multiples of the N above, and A has rank two, $ab \neq 0$. But (164) is non-alternate, while A is alternate and cannot even be congruent to (164).

8. **Reduction to primary components.** Consider first the question of reducing a J-symmetric matrix to primary components. We let $\mathfrak{F}$ be an arbitrary field and let J be an involution defined by

$$A^J = E^{-1}A'E, \tag{165}$$

where $E' = \epsilon E$ is non-singular, $\epsilon = \pm 1$. Suppose that

$$A^J = \delta A, \qquad \delta = \pm 1, \tag{166}$$

and that

$$A = P\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} P^{-1}, \tag{167}$$

where G and H are relatively prime according to our definition. Then $EA^J = A'E = \delta EA$,

$$\delta EP\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} P^{-1} = P'^{-1}\begin{pmatrix} G' & 0 \\ 0 & H' \end{pmatrix} P'E, \tag{168}$$

so that

$$P'EP\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} = \begin{pmatrix} \delta G' & 0 \\ 0 & \delta H' \end{pmatrix} P'EP. \tag{169}$$

By Lemma 12 we have

$$P'EP = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \qquad E_1' = \epsilon E_1, \qquad E_2' = \epsilon E_2. \tag{170}$$

Also $E_1 G = \delta G' E_1$, $E_2 H = \delta H' E_2$, so that we have the first part of the following theorem:

THEOREM 44. *Let $E' = \epsilon E$, $A^J = E^{-1}A'E = \delta A$, where $\epsilon = \pm 1$, $\delta = \pm 1$ and the square matrix A has relatively prime components G and H. Then*

$$A = P\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} P^{-1}, \quad P'EP = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \quad E_1' = \epsilon E_1, \quad E_2' = \epsilon E_2, \tag{171}$$

and the matrices $G^{J_1} \equiv E_1^{-1}G'E_1 = \delta G$, $H^{J_2} \equiv E_2^{-1}H'E_2 = \delta H$. The components G and H are not unique, but if

$$A = P_0\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} P_0^{-1}, \tag{172}$$

then $G_0 = Q_1^{-1}GQ_1$, $H_0 = Q_2^{-1}HQ_2$, *and the replacement of* G *and* H *by similar matrices replaces* E_1 *and* E_2 *by congruent matrices such that*

$$P_0' E P_0 = \begin{pmatrix} Q_1' E_1 Q_1 & 0 \\ 0 & Q_2' E_2 Q_2 \end{pmatrix}. \tag{173}$$

The last part of the above is due to Lemma 12 and

$$P^{-1}P_0\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} = \begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} P^{-1}P_0, \qquad P_0 = P\begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix}. \tag{174}$$

We now let A_0 be a matrix similar to A and such that $A_0^J = \epsilon A_0$. Then

$$A_0 = R\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} R^{-1}, \qquad R'ER = \begin{pmatrix} L_1 & 0 \\ 0 & L_2 \end{pmatrix}. \tag{175}$$

If $A_0 = DAD^{-1}$, where $D^J D = I$, then

$$DP\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}(DP)^{-1} = R\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} R^{-1}, \tag{176}$$

so that

$$DP = R\begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}, \qquad G = C_1^{-1}G_0C_1, \qquad H_0 = C_2^{-1}H_0C_2. \tag{177}$$

Moreover $E^{-1}D'ED = I$,

$$(DP)'E(DP) = P'EP = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} = \begin{pmatrix} C_1' L_1 C_1 & 0 \\ 0 & C_2' L_2 C_2 \end{pmatrix}. \tag{178}$$

Hence necessarily L_1 and E_1 are congruent; L_2 and E_2 are congruent. Hence we may replace G_0 and H_0 by similar matrices and take $E_1 = L_1$, $E_2 = L_2$. We may now prove the following theorem:

THEOREM 45. *Let* $A_0^J = \delta A_0$ *be similar to* A *of Theorem 44. Then* A_0 *is* J*-orthogonally equivalent to* A *only if*

$$A_0 = R\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} R^{-1}, \qquad R'ER = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \tag{179}$$

so that $G_0^{J_1} = \delta G_0$, $H_0^{J_2} = \delta H_0$. *Moreover* A_0 *is* J*-orthogonally equivalent to* A *if and only if* G_0 *is* J_1*-orthogonally equivalent to* G, *and* H_0 *is* J_2*-orthogonally equivalent to* H.

This theorem reduces the problem of J-orthogonal equivalence to the similar problem for primary matrices. For proof we need only take $L_1=E_1$, $L_2=E_2$ above and obtain $C_1' E_1C_1=E_1$, $C_1^JC_1=I_1$ and similarly $C_2^JC_2=I_2$. Thus $G_0=C_1GC_1^J$, $H_0=C_2HC_2^J$ as desired.

We apply Theorem 45 to our case of orthogonal equivalence in $\mathfrak{F}$ of characteristic two. Suppose that $A=A'$ has components G and H. Then

$$P'IP = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}. \tag{180}$$

One of E_1 and E_2 must be definite by Theorem 7. Assume that E_1 is definite so that we may take $E_1=I_1$ to be an identity matrix. Then two cases arise. In the first case E_2 is definite; we may take $E_2=I_2$ and have $P'P=I$; P is an orthogonal matrix and A is orthogonally equivalent to

$$\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}$$

where both G and H are symmetric. Moreover

$$A_0 = P_0\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix}P_0^{-1}, \qquad P_0^{-1}P_0 = I,$$

is orthogonally equivalent to A if and only if G and G_0 are orthogonally equivalent, and H and H_0 are orthogonally equivalent.

We next consider the only remaining case, that where E_2 is an alternate matrix. Then H has even order $2s$, and we may take

$$E_2 = \begin{pmatrix} 0 & I_s \\ I_s & 0 \end{pmatrix}.$$

The matrix H is J_2-symmetric by Theorem 42, and

$$H^{J_2} = E_2^{-1}H'E_2.$$

However G is symmetric. Finally if A_0 is similar to A, then necessarily

$$A_0 = P_0\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix}P_0^{-1}, \qquad P_0^{-1}P_0 = \begin{pmatrix} I_1 & 0 \\ 0 & E_2 \end{pmatrix},$$

if A_0 is to be orthogonally equivalent to A. We apply this result to prove the following theorem:

THEOREM 46. *There exist symmetric matrices A which have symmetric relatively prime components G, H and are such that A is not orthogonally equivalent to any direct sum*

$$\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} \tag{181}$$

for G_0 similar to G, H_0 similar to H.

For let us construct any symmetric matrix G^* whose characteristic function is prime to that of H, where

$$H = \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix}, \qquad E_2 = \begin{pmatrix} 0 & I_s \\ I_s & 0 \end{pmatrix}.$$

Then $H^{J_2} = E_2^{-1}H'E_2 = H$ by direct computation. Also

$$\begin{pmatrix} I_{n-2s} & 0 \\ 0 & E_2 \end{pmatrix} = P'P$$

is definite. Hence

$$A = P\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}P^{-1}$$

is symmetric and has G and H as components. But if A were orthogonally equivalent to (181), we could apply Theorem 44 to see that E_2 is congruent to the identity matrix I_{2s}. But this is impossible since E_2 is alternate.

We now use the form of (180) to prove the theorem:

THEOREM 47. *A matrix A with elements in a field $\mathfrak{F}$ of characteristic two is similar to a symmetric matrix if and only if the minimum function of A is not a product of distinct inseparable irreducible polynomials.*

For every symmetric A is similar to a direct sum of its relatively prime primary components. By (180) and Theorem 45 one of these primary components B_i is similar to a symmetric matrix. But then Theorem 41 implies that the minimum function of B_i, which is a factor of that of A, is not an inseparable irreducible polynomial. Conversely let A have minimum function

$$f(x) = g(x) \cdot h(x),$$

where $g(x)$ and $h(x)$ are relatively prime, $h(x)$ is the product of all irreducible inseparable factors of $f(x)$ whose second power does not divide $f(x)$. By Theorem 41 corresponding to each distinct power $[g_i(x)]^{e_i}$ of an irreducible polynomial occurring in the factorization of $g(x)$ into such factors there corresponds a symmetric matrix G_i which is a primary component of A. Their direct sum is a symmetric component G of A. By Theorem 36 correspond-

* In particular we might take G to define a symmetric separable field, S nilpotent.

ing to each irreducible factor $h_i(x)$ of $h(x)$ there is an indecomposable matrix H_i with $h_i(x)$ as characteristic function and such that E_i is alternate, $E_i^{-1}H_i'E_i = H_i$. The direct sum of the H_i is a matrix H which is a component of A such that A is similar to

$$\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}, \qquad G = G', \qquad E^{-1}H'E = E,$$

where E is the alternate matrix which is the direct sum of the E_i. But

$$\begin{pmatrix} I & 0 \\ 0 & E \end{pmatrix} = P'P; \qquad P\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}P^{-1}$$

is symmetric by Theorem 44 and is similar to A.

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

GENERALIZED INTEGRALS AND DIFFERENTIAL EQUATIONS*

BY
HANS LEWY

Introduction. The idea of the following considerations can best be explained in the simplest case of an integral

$$\int a(x, f(x))db(x, f(x))$$

in which a and b are continuously differentiable functions of two variables, $f(x)$ a continuously differentiable function. The routine estimate of $\int adb$ gives bounds depending either on $\int |df(x)|$ or on the maximum of $|df/dx|$ in the interval of integration. It is, however, possible, as shown in Theorem 1, to give a bound that is entirely independent of the derivative of $f(x)$, and, consequently, to define, by a limiting process, $\int adb$, even in the case where $f(x)$ not only has no derivative, but is no longer continuous, provided $f(x)$ belongs to Baire's first class. The same observation holds for a great number of functionals of $f(x)$ whose construction depends on the derivative of $f(x)$, but for which bounds can be found nevertheless without reference to df/dx. In this paper we are concerned mainly with ordinary differential equations (Theorems 2–3′) and systems of hyperbolic equations in two independent variables (Theorem 6 and corollaries) whose treatment is based on a detailed study of the double integral (33).

1. **Simple integrals.** The theory of the Lebesgue-Stieltjes integral contains the following statement: If in a closed interval J the function $\beta_0(x)$ is monotone and the sequence of continuous functions $\alpha_\mu(x)$ is uniformly bounded and tends to a limit function $\alpha_0(x)$, then the Stieltjes integral $\int\alpha_\mu(x)d\beta_0(x)$ tends to the Lebesgue-Stieltjes integral $\int\alpha_0(x)d\beta_0(x)$ as $\mu\rightarrow 0$. If, furthermore, a sequence of functions $\beta_\mu(x)$ of bounded variation tends to $\beta_0(x)$ as $\mu\rightarrow\infty$ so that the total variation of the difference $\beta_\mu(x)-\beta_0(x)$ tends to zero, then

$$\limsup \left| \int \alpha_\mu(x)d\beta_\mu(x) - \int \alpha_0(x)d\beta_0(x) \right|$$
$$\leqq \limsup \left\{\max |\alpha_\mu(x)| \cdot \int |d(\beta_\mu - \beta_0)|\right\} \rightarrow 0,$$

* Presented to the Society, November 28, 1936; received by the editors May 20, 1936, and May 24, 1937.

which relation leads to the following lemma:

LEMMA 1. *If in the interval $X_0 \leqq x \leqq X_1$ the continuous functions $\alpha_\mu(x)$, $\mu=1, 2, \cdots$, are uniformly bounded and tend to $\alpha_0(x)$ as $\mu\rightarrow\infty$, and if the functions $\beta_\mu(x)$ of bounded variation tend to a function $\beta_0(x)$ of bounded variation while the total variation of $\beta_\mu(x)-\beta_0(x)$ tends to zero as $\mu\rightarrow\infty$, then*

$$\lim_{\mu\rightarrow\infty}\int_{X_0}^{x}\alpha_\mu(x)d\beta_\mu(x) \text{ exists and } = \int_{X_0}^{x}\alpha_0(x)d\beta_0(x),$$

where the integral on the right is the Lebesgue-Stieltjes integral.

We proceed to introduce a new notion of integral which is essentially different from the Lebesgue-Stieltjes integral and is based upon the following theorem:

THEOREM 1. *In the interval J $(X_0 \leqq x \leqq X_1)$ there are given a function $f(x)$ satisfying the inequality $|f(x)-f(X_0)|<F$, a sequence $\{f_\mu(x)\}$ of continuously differentiable functions with $|f_\mu(x)-f(X_0)|<F$ such that $f_\mu(x)\rightarrow f(x)$ as $\mu\rightarrow\infty$, and n continuous functions $g_1(x), \cdots, g_n(x)$ with bounded total variations $T[g_1], \cdots, T[g_n]$. Denote by $f^0, g_1^0, \cdots, g_n^0$ the values of $f(X_0), g_1(X_0), \cdots, g_n(X_0)$ and by $\gamma_1, \cdots, \gamma_n$ upper bounds of $|g_1(x)-g_1^0|, \cdots, |g_n(x)-g_n^0|$ in J. Let $\{g_{1\mu}(x)\}, \cdots, \{g_{n\mu}(x)\}$ be n sequences of continuous functions of bounded variation, tending to $g_1(x), \cdots, g_n(x)$, respectively, as $\mu\rightarrow\infty$ while the total variations of the differences $T[g_{1\mu}-g_1], \cdots, T[g_{n\mu}-g_n]$ tend to zero and $|g_{i\mu}(x)-g_i^0|<\gamma_i$ for all x in J, $i=1, 2, \cdots, n$ and $\mu=1, 2, \cdots$. Suppose that in the $(n+2)$-dimensional domain*

$$D: \quad \begin{array}{l} |z-f^0| \leqq F,\ X_0 \leqq x \leqq X_1, \\ |y_1-g_1^0| \leqq \gamma_1, \cdots, |y_n-g_n^0| \leqq \gamma_n \end{array}$$

the functions $a(z, x, y_1, \cdots, y_n)$ and $b(z, x, y_1, \cdots, y_n)$ are continuously differentiable. Then the Stieltjes integral

$$(1)\quad S_\mu(x) = \int_{X_0}^{x} a(f_\mu(x), x, g_{1\mu}(x), \cdots, g_{n\mu}(x))db(f_\mu(x), x, g_{1\mu}(x), \cdots, g_{n\mu}(x))$$

tends to a limit $L(x)$ as $\mu\rightarrow\infty$.

Remarks. It is clear that any function $f(x)$ with $|f(x)-f^0|<F$ which is a limit of continuous functions may be considered as a limit of continuously differentiable functions $f_\mu(x)$ with $|f_\mu(x)-f^0|<F$. Moreover, the limit $L(x)$ is independent of the approximating sequences $f_\mu(x)$ and $g_{i\mu}(x)$. For the statement of Theorem 1 implies the existence of a limit no matter which sequences are used and hence any two sequences may be considered as subsequences of

a third one containing both. Consequently we may state the following as a definition:

DEFINITION.

$$L(x) = \int_{X_0}^{x} a(f(x), x, g_1(x), \cdots, g_n(x))db(f(x), x, g_1(x), \cdots, g_n(x)).$$

In order to prove Theorem 1, we first assume that $b(z, x, y_1, \cdots, y_n)$ has continuous derivatives of second order of the mixed type. We determine a function $A(z, x, y_1, \cdots, y_n)$ in D as the solution of the differential equation

$$\frac{\partial A}{\partial z} = a\frac{\partial b}{\partial z}$$

with the initial condition $A(f^0, x, y_1, \cdots, y_n) = 0$. We obtain

$$\begin{aligned} &A(z, x, y_1, \cdots, y_n) \\ &\quad = \int_{f^0}^{z} a(z', x, y_1, \cdots, y_n)b_z(z', x, y_1, \cdots, y_n)dz', \end{aligned}$$

where we may differentiate with respect to $x, y_1, \cdots, y_n$ under the integral sign. Thus we find

$$\begin{aligned} adb &= ab_z dz + ab_x dx + \sum_{i=1}^{n} ab_{y_i} dy_i \\ &= dA + (ab_x - A_x)dx + \sum (ab_{y_i} - A_{y_i})dy_i, \end{aligned}$$

$$\begin{aligned} A_x(z, x, y_1, \cdots, y_n) &= \int_{f^0}^{z} (ab_z)_x dz' \\ &= \int_{f^0}^{z} [(ab_x)_z + (a_x b_z - a_z b_x)]dz' \\ &= a(z, x, y_1, \cdots)b_x(z, x, y_1, \cdots) \\ &\quad - a(f^0, x, y_1, \cdots)b_x(f^0, x, y_1, \cdots) + \int_{f^0}^{z} (a_x b_z - a_z b_x)dz' \end{aligned}$$

and

$$\begin{aligned} A_{y_i}(z, x, y_1, \cdots, y_n) &= a(z, x, y_1, \cdots)b_{y_i}(z, x, y_1, \cdots) \\ &\quad - a(f^0, x, y_1, \cdots)b_{y_i}(f^0, x, y_1, \cdots) \\ &\quad + \int_{f^0}^{z} (a_{y_i} b_z - a_z b_{y_i})dz'. \end{aligned}$$

Hence

$$S_\mu(x) \equiv \int_{X_0}^{x} a(f_\mu(x'), x', g_{i\mu}(x'))db(f_\mu(x'), x', g_{i\mu}(x'))$$

$$= A(f_\mu(x), x, g_{i\mu}(x)) + \int_{X_0}^{x} a(f^0, x', g_{i\mu}(x'))b_x(f^0, x', g_{i\mu}(x'))dx$$

$$(2) \qquad + \sum_{i=1}^{n}\int_{X_0}^{x} a(f^0, x', g_{1\mu}(x'), \cdots)b_{y_i}(f^0, x', g_{1\mu}(x'), \cdots)dg_{i\mu}(x')$$

$$+ \int_{X_0}^{x} dx' \int_{f^0}^{f_\mu(x')} [a_z(z', x', g_{1\mu}(x'), \cdots)b_x(z', x', g_{1\mu}(x'), \cdots) - a_x b_z(\cdots)]dz'$$

$$+ \sum_i \int_{X_0}^{x} dg_{i\mu}(x') \int_{f^0}^{f_\mu(x')} [a_z(z', x', g_{1\mu}(x'), \cdots)b_{y_i} - a_{y_i}b_z]dz'.$$

This formula, derived under the assumption that b has continuous second derivatives of mixed type, still holds under the conditions of Theorem 1. For any b which is continuously differentiable in D may be uniformly approximated by a polynomial such that its first derivatives uniformly approximate those of b. Introducing the approximations instead of b into (2) and passing to the limit we obtain again the formula (2) as both sides of (2) involve only first derivatives of b and the passage to the limit under the integral signs is legitimate in view of the uniform convergence of the derivatives of the polynomials to those of b.

In the right-hand member of (2) we can effect the passage $\mu \to \infty$ by simply cancelling all reference to μ. This may be seen as follows. An integral $\int_{f^0}^{z} a_z(z', x, y_1, \cdots, y_n)b_{y_i}dz'$, for instance, is continuous in $z, x, y_1, \cdots, y_n$. Thus $\int_{f^0}^{f_\mu(x)} a(z', x, g_{1\mu}(x), \cdots)b_{y_i}dz'$ is a continuous function of x, bounded as $\mu \to \infty$, and converging as $\mu \to \infty$. Now the convergence of

$$\int_{X_0}^{x} dg_{i\mu}(x') \int_{f^0}^{f_\mu(x)} a_z(z', x', g_{1\mu}(x'), \cdots)b_{y_i}dz'$$

follows from Lemma 1.

Thus Theorem 1 is proved.

From (2) we have the following estimate:

$$|L(x)| \leqq MN\left\{|f(x) - f^0| + \sum_{i=1}^{n} T_{X_0}^{x}[g_i] + |x - X_0|\right\}$$

$$+ 2N^2F\left(\sum_i T_{X_0}^{x}[g_i] + |x - X_0|\right),$$

where M is an upper bound for $|a|$ and $|b|$, N an upper bound for the moduli of the first derivatives of a and b in D.

Remark. We have, for instance, for every admissible $f(x)$

$$L(x) \equiv \int_0^x f(x')df(x') = \tfrac{1}{2}(f^2(x) - f^2(0)),$$

which leads to $L(1)=\frac{1}{2}$ for $f(x)=0$ if $0 \leqq x < 1, f(1)=1$. The Lebesgue-Stieltjes integral, however, would be 1.

2. **Ordinary differential equations.** We may now prove the following theorem:

THEOREM 2. *Let the functions $f(x)$, $g_1(x), \cdots, g_n(x)$ be continuously differentiable in $0 \leqq x \leqq X$ and $f(0)=f^0$, $g_1(0)= \cdots =g_n(0)=0$. Denote by F an upper bound of $|f(x)|$ and by $G_1, G_2, \cdots, G_n$ the total variations of $g_1(x)$, $g_2(x), \cdots, g_n(x)$ in $[0, X]$. Suppose, for $\epsilon > 0$, that in the domain*

$$D_{3,\epsilon}: \quad |z| \leqq F, \quad |y_i| \leqq G_i, \quad |u| < \epsilon + 3FM_0 + \sum_{i=1}^{n} (e^{2FN} - e^{FN} + M_i e^{FN})G_i$$

the functions $a_0, a_1, \cdots, a_n$ are continuous functions of z, y_i, u satisfying the inequalities $|a_k| \leqq M_k$, $k=0, 1, \cdots, n$, and that a_0 has continuous derivatives of first order, bounded in absolute value by N. Then the solution $u(x)$ of the equation

$$\text{(E)} \qquad \begin{aligned} du(x) &= a_0(f(x), g_1(x), \cdots, g_n(x), u(x))df(x) \\ &\quad + \sum_{i=1}^{n} a_i(f(x), g_1(x), \cdots, g_n(x), u(x))dg_i(x) \end{aligned}$$

with $u(0)=0$ can be extended over the whole interval $0 \leqq x \leqq X$. It satisfies, moreover, the inequality

$$|u(x)| \leqq 2FM_0 + \sum_{i=1}^{n} (e^{2FN} - e^{FN} + M_i e^{FN})G_i.$$

Let us solve the auxiliary partial equation for $A(z, y_1, \cdots, y_n, u)$

$$\text{(3)} \qquad \frac{\partial A}{\partial z} + \frac{\partial A}{\partial u}\cdot a_0 = a_0, \qquad 1 - \frac{\partial A}{\partial u} \neq 0$$

under the initial condition

$$\text{(4)} \qquad A(z, y_i, u) = 0 \text{ for } z = 0.$$

The characteristic equations of (3) are

$$\text{(5)} \qquad dz:du:dA = 1:a_0:a_0.$$

Consider the family C of curves satisfying the differential equation $du/dz=a_0$ and passing through any point P of the domain

$$D_{2,\epsilon}:\qquad |z|\leqq F,\ |y_i|\leqq G_i,\ |u|<\epsilon+2FM_0+\sum_{i=1}^{n}(e^{2FN}-e^{FN}+M_ie^{FN})G_i.$$

Since a_0 has continuous derivatives of first order throughout $D_{2,\epsilon}$ and is bounded by M_0, there exists one and only one curve through an arbitrary point P, and the corresponding value $\bar{u}$ of u for $z=0$ lies within the range

$$|\bar{u}|<\epsilon+3FM_0+\sum_{i=1}^{n}(e^{2FN}-e^{FN}+M_ie^{FN})G_i.$$

Conversely, an arbitrary curve of C is uniquely determined by the quantities $\bar{u}, y_1, \cdots, y_n$, and a point of the curve is determined by giving in addition the corresponding value of z. On writing $u=u(z, y_i, \bar{u})$, we find bounds for the derivatives of u with respect to the arguments $\bar{u}, y_1, \cdots, y_n, z$. We have, in fact,

$$\frac{d}{dz}\frac{\partial u}{\partial\bar{u}}=\frac{\partial a_0}{\partial u}\frac{\partial u}{\partial\bar{u}}\quad\text{with}\quad\frac{\partial u}{\partial\bar{u}}=1\quad\text{for}\quad z=0,$$

hence

$$e^{-N|z|}\leqq\frac{\partial u}{\partial\bar{u}}\leqq e^{N|z|}.$$

Similarly

$$\frac{d}{dz}\frac{\partial u}{\partial y_i}=\frac{\partial a_0}{\partial y_i}+\frac{\partial a_0}{\partial u}\frac{\partial u}{\partial y_i}\quad\text{with}\quad\frac{\partial u}{\partial y_i}=0\quad\text{for}\quad z=0,$$

hence

$$\left|\frac{\partial u}{\partial y_i}\right|\leqq e^{N|z|}-1.$$

On introducing the new variables $z, y_1, \cdots, y_n, u$ instead of $z, y_1, \cdots, y_n, \bar{u}$ we find throughout $D_{2,\epsilon}$

$$\begin{gathered}\frac{\partial(z, y_i, u)}{\partial(z, y_i, \bar{u})}=\frac{\partial u}{\partial\bar{u}},\quad\frac{\partial(z, y_i, \bar{u})}{\partial(z, y_i, u)}=\frac{\partial\bar{u}}{\partial u}=\left(\frac{\partial u}{\partial\bar{u}}\right)^{-1},\quad e^{-|z|N}\leqq\frac{\partial\bar{u}}{\partial u}\leqq e^{|z|N},\\ \frac{\partial\bar{u}}{\partial y_i}=-\frac{\partial u}{\partial y_i}\Big/\frac{\partial u}{\partial\bar{u}},\quad\left|\frac{\partial\bar{u}}{\partial y_i}\right|\leqq e^{|z|N}(e^{|z|N}-1).\end{gathered}\tag{6}$$

Since $\bar{u}$ is constant along any curve of C, we have

$$\frac{\partial \bar{u}}{\partial z} + a_0 \frac{\partial \bar{u}}{\partial u} = 0, \qquad \left| \frac{\partial \bar{u}}{\partial z} \right| \leqq M_0 e^{N|z|}.$$

Now put, throughout $D_{2,\epsilon}$,

$$A(z, y_i, u) = u - \bar{u}. \tag{7}$$

Evidently A satisfies (3) and (4). Furthermore we have

$$\begin{gathered} |A| \leqq |z| M_0, \qquad \left| \frac{\partial A}{\partial y_i} \right| \leqq e^{|z|N}(e^{|z|N} - 1), \\ \left| \frac{\partial A}{\partial z} \right| \leqq M_0 e^{N|z|}, \qquad e^{-|z|N} \leqq \left| 1 - \frac{\partial A}{\partial u} \right| \leqq e^{|z|N}. \end{gathered} \tag{8}$$

Returning to the ordinary differential equation (E), we remark that the conditions of our theorem allow us to write (E) in the form

$$\frac{du}{dx} = \phi(x, u)$$

with $\phi(x, u)$ continuous in the rectangle R determined by

$$0 \leqq x \leqq X, \quad |u| \leqq \frac{\epsilon}{2} + 2FM_0 + H, \quad H = \sum_{i=1}^{n} (e^{2FN} - e^{FN} + M_i e^{FN}) G_i.$$

The fundamental existence theorem for differential equations shows that a solution through $[x=0, u(0)=0]$ always may be continued until it reaches the boundary of R. Hence a solution which cannot be continued across a certain point x_1 with $0 \leqq x_1 < X$ may be assumed to exist for $0 \leqq x \leqq x_1$ and to satisfy the condition

$$|u(x_1)| = \frac{\epsilon}{2} + 2FM_0 + H.$$

Thus our theorem is proved as soon as we show the following property of $u(x)$. If, in the interval $0 \leqq x \leqq x_1$, the solution $u(x)$ of (E) satisfies the inequality

$$|u(x)| \leqq \frac{\epsilon}{2} + 2FM_0 + H,$$

it satisfies the stronger inequality

$$|u(x)| \leqq 2FM_0 + H.$$

Indeed, by (E) we have, since $z = f(x)$, $y_i = g_i(x)$, $u = u(x)$ stay in $D_{2,\epsilon}$,

$$(9)\quad \begin{aligned} d(u - A) &= (a_0 - A_z)df(x) + \sum_{i=1}^{n} (a_i - A_{y_i})dg_i(x) - A_u du \\ &= (a_0 - A_z - A_u a_0)df(x) + \sum_{i=1}^{n} (a_i(1 - A_u) - A_{y_i})dg_i(x), \end{aligned}$$

with $z=f(x)$, $y_i=g_i(x)$ and $u(0)=0$. Hence we conclude from (3) and (8)

$$(10)\qquad | u(x) - A(f(x), g_i(x), u(x)) | \leqq FM_0 + H,$$

and, by (8),

$$(11)\qquad | u(x) | \leqq 2FM_0 + H.$$

Remarks. The function $u(z, y_i, \bar{u})$ is continuous. This may be expressed by the statement: If the quantities z, y_i, $u-A(z, y_i, u)$ tend to limiting values which belong to the domain $|z| \leqq F$, $|y_i| \leqq G_i$, $|u-A(z, y_i, u)| \leqq H+FM_0$, then u itself tends to a limiting value which in absolute value does not exceed $2FM_0+H$.

Any function satisfying a Lipschitz condition of exponent 1 in z, y_i, u satisfies also a Lipschitz condition of exponent 1 in the variables z, y_i, $u-A$. This follows from (8), for we have

$$(12)\qquad | u_1 - A(z, y_i, u_1) - u_2 + A(z, y_i, u_2) | \geqq | u_2 - u_1 | e^{-FN}.$$

THEOREM 3. *Assume that in the interval* $0 \leqq x \leqq X$

(i) *the functions* $f_\mu(x)$, $\mu=1, 2, \cdots$, *are continuously differentiable,* $|f_\mu(x)| < F$, *and* $f_\mu(x)$ *converges to a function* $f(x)$ *as* $\mu \to \infty$;

(ii) *the functions* $g_{1\mu}(x), \cdots, g_{n\mu}(x)$, $\mu=1, 2, \cdots$, *are continuously differentiable,* $g_{i\mu}(x) \to g_i(x)$ *as* $\mu \to \infty$, $i=1, 2, \cdots, n$, *where* $g_i(x)$ *is continuous, and the total variations of the differences* $T[g_{i\mu}-g_i]$ *tend to zero as* $\mu \to \infty$; *furthermore* $g_{i\mu}(0)=0$ *and* $T[g_{i\mu}] \leqq G_i$ *for all* i *and* μ;

(iii) *the functions* $a_0, a_1, \cdots, a_n$ *are defined in*

$$D_{3,\epsilon}:\quad \begin{aligned} &| z | \leqq F, \quad | y_i | \leqq G_i, \quad | u | < \epsilon + 3FM_0 + H, \\ &H = \sum_{i=1}^{n} (e^{2FN} - e^{FN} + M_i e^{FN})G_i, \qquad (\epsilon > 0), \end{aligned}$$

and we have in $D_{3,\epsilon}$

$$(13)\qquad | a_0 | \leqq M_0, \ | a_1 | \leqq M_1, \cdots, \ | a_n | \leqq M_n.$$

Furthermore, $a_0, a_1, \cdots, a_n$ *have continuous derivatives of first order in* $D_{3,\epsilon}$, *and those of* a_0 *are bounded in absolute value by* N *and satisfy a Lipschitz condition of exponent* 1.

Then the solution of

$$(\mathrm{E}_\mu) \quad du_\mu(x) = a_0(f_\mu(x), g_{i\mu}(x), u_\mu(x))df_\mu(x) + \sum_{i=1}^{n} a_i dg_{i\mu}(x), \quad u_\mu(0) \equiv 0,$$

exists for $0 \leqq x \leqq X$, *satisfies*

$$(14) \qquad | u_\mu(x) | \leqq 2FM_0 + H,$$

and tends to a limit function $u(x)$ *as* $\mu \to \infty$. $u(x)$ *is said to be the solution of* (E) *for the initial condition* $u(0) = 0$.

From Theorem 2 we conclude the existence of $u_\mu(x)$ and the inequality (14). The classical statement about uniqueness of the solution of the initial problem may, incidentally, be used to establish the uniqueness of $u_\mu(x)$. On putting

$$B_\mu(x) = u_\mu(x) - A(f_\mu(x), g_{i\mu}(x), u_\mu(x)),$$

we conclude from (9) and (8)

$$(15) \qquad | B_\mu(x_1) - B_\mu(x_2) | \leqq \sum_{i=1}^{n} (M_i e^{FN} + e^{2FN} - e^{FN}) \int_{x_1}^{x_2} | dg_{i\mu}(x) | .$$

Now a theorem by Adams and Clarkson† shows that the total variation between any two points x_1 and x_2, of $g_{i\mu}(x)$ tends uniformly to that of $g_i(x)$ on account of the continuity of $g_i(x)$, and of the assumption (iii) that $T[g_{i\mu} - g_i] \to 0$, $g_{i\mu} \to g_i$. Thus (15) establishes equicontinuity for $B_\mu(x)$, while (14) gives boundedness. Hence, by Ascoli's theorem, we may select a subsequence $B_{\mu'}(x)$ tending uniformly to a function $B^*(x)$. From the remark on page 444 we conclude that also the corresponding subsequence of $u_\mu(x)$, say $u_{\mu'}(x)$, converges to a function $u^*(x)$. $B^*(x)$ satisfies the following integral equation

$$(16) \qquad B^*(x) = \int_0^x \sum_{i=1}^{n} (a_i(1 - A_u) - A_{y_i})dg_i(x) - A(f^0, 0, 0, \cdots)$$

in which the expressions in a_i and A are to be considered as functions of z, y_i, u with $z = f(x)$, $y_i = g_i(x)$, $u - A = B^*(x)$. This follows from Lemma 1 and (9).

Any two subsequences of $B^*(x)$ converge to the same limit. In the opposite case we would have two functions $B^*(x)$ and $B^{**}(x)$, both satisfying (16). In view of (iii) the coefficients $a_0, a_1, \cdots, a_n$ admit of continuous derivatives with respect to z, y_i, u, whence also with respect to z, y_i, $u - A$, and thus satisfy a Lipschitz condition of exponent 1 in these variables, in the closed domain $|z| \leqq F$, $|y_i| \leqq G_i$, $|u - A| \leqq FM_0 + H$. On account of (3) and (4), the

† C. R. Adams and J. A. Clarkson, *On convergence in variation*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 413–417.

derivatives $A_{y'}$ and A_u satisfy a Lipschitz condition with respect to z, y_i, u, whence with respect to z, y_i, $u-A$. Therefore we conclude from (16)

$$(17)\qquad |B^*(x) - B^{**}(x)| \leqq K\sum_i \int_0^x |B^*(x) - B^{**}(x)||dg_i(x)|,$$

where K is a certain constant. On iterating (17) m times we easily find

$$(18)\quad \max_{0\leqq x\leqq X} |B^*(x) - B^{**}(x)| \leqq K^m \max_{0\leqq x\leqq X} |B^*(x) - B^{**}(x)| \left(\sum_{i=1}^n G_i\right)^m \Big/ m!$$

which gives $B^*(x)-B^{**}(x)=0$ as $M\rightarrow\infty$.

Now the uniqueness of $B^*(x)$ implies, by the remark on page 444, the uniqueness of $u^*(x)$, which in turn justifies defining $u(x)=u^*(x)$ as the solution of (E) for the initial condition $u(0)=0$.

Remark. The assumptions of Theorem 3 state bounds for the functions $a_0, a_1, \cdots, a_n(z, y_1, \cdots, y_n, u)$ holding in a domain that depends on these same bounds. One may ask for a formulation of the theorem such that a statement results for *any* functions $a_0, a_1, \cdots, a_n$, defined in an *arbitrary* neighborhood of the origin. Therefore we observe that there always exists a sub-neighborhood of the form $D_{3,\epsilon}$, provided the constants $F, G_1, \cdots, G_n$, ϵ can be decreased sufficiently. Since the functions $g_{1\mu}(x), \cdots, g_{n\mu}(x)$ are continuous, their total variations are also continuous and may be shown† to converge *uniformly* to those of $g_1(x), \cdots, g_n(x)$. Thus, omitting at most a finite number of values of μ and taking the upper end of the x-interval sufficiently small, makes it possible to choose $G_1, G_2, \cdots, G_n$ arbitrarily small. Whence we conclude the following theorem:

THEOREM 3′. *Suppose that, in a neighborhood of the origin of a $(z, y_1, \cdots, y_n, u)$-space the functions $a_0, a_1, \cdots, a_n$ are continuously differentiable and the derivatives of a_0 satisfy a Lipschitz condition of exponent* 1. *Assume that n continuously differentiable functions $g_{1\mu}(x), \cdots, g_{n\mu}(x)$ are defined in an interval I $(0\leqq x\leqq X)$, that they tend to continuous functions $g_1(x), \cdots, g_n(x)$, and the total variations $T[g_{i\mu}-g_i]\rightarrow 0$ as $\mu\rightarrow\infty$, and that $g_{i\mu}(0)=0$. Assume furthermore that continuously differentiable functions $f_\mu(x)$ converge to a function $f(x)$ in I and that for all μ we have $|f_\mu(x)|<F$. Then the solution $u_\mu(x)$ of* (E_μ) *exists in a sufficiently small interval $0\leqq x\leqq X'$ which does not depend on μ, and converges there to a limit function $u(x)$ provided F is sufficiently small.*

The existence Theorems 3 and 3′ can be supplemented by a study of the manner in which the solution $u(x)$ depends on a parameter α on which the known functions in (E) may be supposed to depend. Usual methods of proving

† See Adams and Clarkson, loc. cit.

the continuity of $u(x)$, considered as a function of x and α, from that of the known functions could be carried through with only slight modifications.

Instead of (E) a system of differential equations of the form

$$(19)\qquad du_h(x) = a_{0h}(f(x), g_i(x), u_l(x))df(x) + \sum_{i=1}^{n} a_{ih}dg_i(x), \qquad u_h(0) = 0,$$

where $h=1, 2, \cdots, m$, may be studied, and the existence of a solution $u_h(x)$, $h=1, 2, \cdots, m$, can be concluded by a method analogous to that used in the proofs of Theorems 2 and 3. In view of the similarity of the procedure we shall not carry out these generalizations.

3. **Double integrals.** We denote by $T[\alpha, \beta]$ a triangle bounded by the line $\alpha=\beta$ of an (α, β)-plane and the parallels to the axes through (α, β). Similarly $t[f, g]$ designates a triangle of an (f, g)-plane, bounded by $f=g$ and the parallels to the f- and g-axes through (f, g). The elements of area $d\alpha d\beta$ and $dfdg$ are to be counted positive. By $f(\alpha)$ and $g(\beta)$ we understand continuously differentiable mappings of the α-axis on the f-axis and of the β-axis on the g-axis, which are, but for the elements used, identical with each other; $f(\alpha)=g(\beta)$ if $\alpha=\beta$. The range of the four variables α, β, f, g is the domain D with origin as center

$$(D)\qquad |\alpha| \leqq \omega, \quad |\beta| \leqq \omega, \quad |f| \leqq \omega, \quad |g| \leqq \omega,$$

and the function $f(\alpha)$ (and consequently $g(\beta)$) is such that for α and β satisfying (D) the point $(\alpha, f(\alpha), \beta, g(\beta))$ belongs to D. Furthermore there are defined in D three functions a, b, c of α, f, β, g having continuous derivatives up to the fourth order.

We introduce three functions X, Y, Z in D by the relations

$$(20)\qquad X_{fg} = ca_f b_g,$$

$$(21)\qquad Y_f = ca_f b_\beta, \qquad Z_g = ca_\alpha b_g,$$

and the initial conditions

$$\left.\begin{array}{ll} (22) & X = X_f = X_g = 0, \\ & Y = 0, \\ (23) & Z = 0, \end{array}\right\} \text{ if } f = g.$$

We find

$$X(\alpha, f, \beta, g) = \iint_{t[f,g]} ca_f b_g(\alpha, f', \beta, g')df'dg'.$$

Here, as in all integrals that follow, care has been taken to indicate the argu-

ments of the integrand at least in one of the factors of the integrand, to denote the variables of integration by a prime and to denote by subscripts the partial derivatives, while we reserve the symbols $d/d\alpha$ and $d/d\beta$ for total derivatives with respect to these variables.

We are going to study the function

$$
\begin{aligned}
I(\alpha, f, \beta, g) = X(\alpha, f, \beta, g) &- \int_\beta^\alpha (X_\alpha(\alpha', f(\alpha'), \beta, g) - Z)d\alpha' \\
&+ \int_\beta^\alpha (X_\beta(\alpha, f, \beta', g(\beta')) - Y)d\beta' \qquad (24)\\
&+ \iint_{T[\alpha,\beta]} (-X_{\alpha\beta}+Z_\beta+Y_\alpha-ca_\alpha b_\beta(\alpha', f(\alpha'),\beta', g(\beta')))d\alpha' d\beta'.
\end{aligned}
$$

In order to abbreviate as much as possible, we write $A \sim B$ if $A-B$ is expressible as a polynomial in a, b, c, their first partial derivatives and their second partial derivatives of the type $\partial^2/\partial\alpha\partial\beta$, $\partial^2/\partial\alpha\partial g$, $\partial^2/\partial\beta\partial f$, $\partial^2/\partial f\partial g$. We write $A \cong B$ if $A-B$ is expressible as an integral over a function which itself is ~ 0.

Thus we find

$$
I_f = X_f + \int_\beta^\alpha (X_{\beta f}(\alpha, f, \beta', g(\beta')) - Y_f)d\beta', \qquad (25)
$$

$$
I_g = X_g - \int_\beta^\alpha (X_{\alpha g}(\alpha', f(\alpha'), \beta, g) - Z_g)d\alpha', \qquad (26)
$$

$$
\begin{aligned}
I_\alpha = X_\alpha(\alpha, f, \beta, g) &- X_\alpha(\alpha, f(\alpha), \beta, g) + Z(\alpha, f(\alpha), \beta, g) \\
&+ X_\beta(\alpha, f, \alpha, f(\alpha)) - Y(\alpha, f, \alpha, f(\alpha)) \\
&+ \int_\beta^\alpha (- X_{\alpha\beta} + Z_\beta + Y_\alpha - ca_\alpha b_\beta(\alpha, f(\alpha), \beta', g(\beta')))d\beta', \qquad (27)\\
&+ \int_\beta^\alpha (X_{\alpha\beta}(\alpha, f, \beta', g(\beta')) - Y_\alpha)d\beta',
\end{aligned}
$$

$$
\begin{aligned}
I_\beta = X_\beta(\alpha, f, \beta, g) &- X_\beta(\alpha, f, \beta, g(\beta)) + Y(\alpha, f, \beta, g(\beta)) \\
&+ X_\alpha(\beta, g(\beta), \beta, g) - Z(\beta, g(\beta), \beta, g) \\
&- \int_\beta^\alpha (X_{\alpha\beta}(\alpha', f(\alpha'), \beta, g) - Z_\beta)d\alpha' \qquad (28)\\
&- \int_\beta^\alpha (- X_{\alpha\beta} + Z_\beta + Y_\alpha - ca_\alpha b_\beta(\alpha', f(\alpha'), \beta, g(\beta)))d\alpha',
\end{aligned}
$$

$$
(29)\qquad
\begin{aligned}
I_{fg} &= ca_f b_g,\\
I_{f\beta} &= X_{f\beta}(\alpha, f, \beta, g) - X_{f\beta}(\alpha, f, \beta, g(\beta)) + Y_f(\alpha, f, \beta, g(\beta)),\\
I_{g\alpha} &= X_{g\alpha}(\alpha, f, \beta, g) - X_{g\alpha}(\alpha, f(\alpha), \beta, g) + Z_g(\alpha, f(\alpha), \beta, g),\\
I_{\alpha\beta} &= X_{\alpha\beta}(\alpha, f, \beta, g) - X_{\alpha\beta}(\alpha, f(\alpha), \beta, g) + Z_\beta(\alpha, f(\alpha), \beta, g)\\
&\quad - Z_\beta(\alpha, f(\alpha), \beta, g(\beta)) + X_{\alpha\beta}(\alpha, f(\alpha), \beta, g(\beta)) - X_{\alpha\beta}(\alpha, f, \beta, g(\beta))\\
&\quad + Y_\alpha(\alpha, f, \beta, g(\beta)) - Y_\alpha(\alpha, f(\alpha), \beta, g(\beta)) + ca_\alpha b_\beta(\alpha, f(\alpha), \beta, g(\beta)).
\end{aligned}
$$

Hence

$$
\begin{aligned}
I_{f\beta} &= \int_{g(\beta)}^{g} (X_{f\beta g} - Y_{fg}(\alpha, f, \beta, g'))dg' + Y_f(\alpha, f, \beta, g)\\
&= \int_{g(\beta)}^{g} [(ca_f b_g)_\beta - (ca_f b_\beta)_g]dg' + ca_f b_\beta\\
&= \int_{g(\beta)}^{g} [(ca_f)_\beta b_g - (ca_f)_g b_\beta]dg' + ca_f b_\beta,
\end{aligned}
$$

$$(30)\qquad I_{f\beta} \cong ca_f b_\beta.$$

Similarly

$$(31)\qquad I_{g\alpha} \cong ca_\alpha b_g.$$

Moreover,

$$
\begin{aligned}
&X_{\alpha\beta}(\alpha, f, \beta, g) - X_{\alpha\beta}(\alpha, f(\alpha), \beta, g) + X_{\alpha\beta}(\alpha, f(\alpha), \beta, g(\beta))\\
&\qquad - X_{\alpha\beta}(\alpha, f, \beta, g(\beta)) = \int_{f(\alpha)}^{f}\int_{g(\beta)}^{g} X_{\alpha\beta fg}(\alpha, f', \beta, g')df'dg'.
\end{aligned}
$$

But

$$
\begin{aligned}
X_{\alpha\beta fg} &= (ca_f b_g)_{\alpha\beta} = (c_\alpha a_f b_g + ca_{f\alpha} b_g + ca_f b_{g\alpha})_\beta\\
&\sim c_\alpha a_f b_{g\beta} + c_\beta a_{\alpha f} b_g + ca_{f\alpha\beta} b_g + ca_{f\alpha} b_{g\beta} + ca_f b_{g\alpha\beta}\\
&\sim (c_\alpha a_f b_\beta)_g + (c_\beta a_\alpha b_g)_f + (ca_{\alpha\beta} b_g)_f + (ca_f b_{\alpha\beta})_g + ca_{\alpha f} b_{\beta g},\\
ca_{\alpha f} b_{\beta g} &= (ca_\alpha b_{\beta g})_f - c_f a_\alpha b_{\beta g} - ca_\alpha b_{g\beta f}\\
&\sim (ca_\alpha b_\beta)_{gf} - [(ca_\alpha)_g b_\beta]_f - (c_f a_\alpha b_\beta)_g - (ca_\alpha b_{\beta f})_g.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\int_{f(\alpha)}^{f}\int_{g(\beta)}^{g} X_{\alpha\beta fg}(\alpha, f', \beta, g')df'dg' &\cong ca_\alpha b_\beta(\alpha, f, \beta, g) - ca_\alpha b_\beta(\alpha, f(\alpha), \beta, g)\\
&\quad + ca_\alpha b_\beta(\alpha, f(\alpha), \beta, g(\beta)) - ca_\alpha b_\beta(\alpha, f, \beta, g(\beta)).
\end{aligned}
$$

Furthermore,

$$\begin{aligned}Z_\beta(\alpha, f(\alpha), \beta, g) - Z_\beta(\alpha, f(\alpha), \beta, g(\beta)) &= \int_{g(\beta)}^{g} Z_{\beta g}(\alpha, f(\alpha), \beta, g')dg' \\ &= \int_{g(\beta)}^{g} [ca_\alpha b_g(\alpha, f(\alpha), \beta, g')]_\beta dg' \\ &\cong \int_{g(\beta)}^{g} [ca_\alpha b_\beta(\alpha, f(\alpha), \beta, g')]_g dg' \\ &\cong ca_\alpha b_\beta(\alpha, f(\alpha), \beta, g) - ca_\alpha b_\beta(\alpha, f(\alpha), \beta, g(\beta)).\end{aligned}$$

Similarly,

$$Y_\alpha(\alpha, f, \beta, g(\beta)) - Y_\alpha(\alpha, f(\alpha), \beta, g(\beta)) \cong ca_\alpha b_\beta(\alpha, f, \beta, g(\beta)) - ca_\alpha b_\beta(\alpha, f(\alpha), \beta, g(\beta)).$$

Finally

$$I_{\alpha\beta}(\alpha, f, \beta, g) \cong ca_\alpha b_\beta(\alpha, f, \beta, g). \tag{32}$$

In view of (29), we find

$$\begin{aligned}\frac{d^2 I}{d\alpha d\beta}(\alpha, f(\alpha), \beta, g(\beta)) &= ca_\alpha b_\beta + ca_f b_\beta \frac{df(\alpha)}{d\alpha} + ca_\alpha b_g \frac{dg(\beta)}{d\beta} + ca_f b_g \frac{df(\alpha)}{d\alpha}\frac{dg(\beta)}{d\beta} \\ &= c\,\frac{da(\alpha, f(\alpha), \beta, g(\beta))}{d\alpha}\,\frac{db(\alpha, f(\alpha), \beta, g(\beta))}{d\beta}\end{aligned}$$

and

$$I(\alpha, f(\alpha), \beta, g(\beta)) = 0, \qquad \frac{dI}{d\alpha} = 0, \qquad \frac{dI}{d\beta} = 0,$$

for $\alpha = \beta$. Thus

$$I(\alpha, f(\alpha), \beta, g(\beta)) = -\iint_{T[\alpha,\beta]} c\,\frac{da}{d\alpha'}\,\frac{db}{d\beta'}(\alpha', f(\alpha'), \beta', g(\beta'))d\alpha' d\beta'. \tag{33}$$

In order to transform $I_f(\alpha, f, \beta, g)$ we calculate

$$\begin{aligned}\int_\beta^\alpha X_{\beta f}(\alpha, f, \beta', g(\beta'))d\beta' &= \int_\beta^\alpha d\beta'\left[\int_f^{g(\beta')} X_{\beta f g}(\alpha, f, \beta', g')dg' + X_{\beta f}(\alpha, f, \beta', f)\right] \\ &= \int_\beta^\alpha d\beta' \int_f^{g(\beta')} (ca_f b_g(\alpha, f, \beta', g'))_\beta dg' \\ &\cong \int_\beta^\alpha d\beta' \int_f^{g(\beta')} (ca_f b_\beta(\alpha, f, \beta', g'))_g dg' \\ &\cong \int_\beta^\alpha d\beta' [ca_f b_\beta(\alpha, f, \beta', g(\beta')) - ca_f b_\beta(\alpha, f, \beta', f)]\end{aligned}$$

since, by (22), $X_{\beta f}(\alpha, f, \beta', f)$ vanishes.

Consequently

$$I_f(\alpha, f, \beta, g) \cong 0. \tag{34}$$

Similarly

$$I_g(\alpha, f, \beta, g) \cong 0. \tag{35}$$

Also we have

$$\begin{aligned} X_\alpha(\alpha, f, \beta, g) &= -\iint_{t[f,g]} (ca_f b_g(\alpha, f', \beta, g'))_\alpha df'dg' \\ &\cong -\iint_{t[f,g]} (ca_\alpha b_g(\alpha, f', \beta, g'))_f df'dg' \\ &\cong -\int ca_\alpha b_g(\alpha, f', \beta, g')dg', \end{aligned}$$

where the integration is to be extended over the boundary of $t[f, g]$; whence

$$X_\alpha(\alpha, f, \beta, g) \cong 0, \qquad X_\beta(\alpha, f, \beta, g) \cong 0.$$

Thus

$$I_\alpha(\alpha, f, \alpha, g) \cong 0, \tag{36}$$

as obviously $Y \cong 0$, $Z \cong 0$.

On writing

$$I_\alpha(\alpha, f, \beta, g) = I_\alpha(\alpha, f, \alpha, g) + \int_\alpha^\beta I_{\alpha\beta}(\alpha, f, \beta', g)d\beta',$$

we find by (36) and (32)

$$I_\alpha(\alpha, f, \beta, g) \cong 0. \tag{37}$$

Similarly,

$$I_\beta(\alpha, f, \beta, g) \cong 0. \tag{38}$$

Finally

$$\begin{aligned} &I(\alpha, f, \beta, g) = I(\alpha, f, \alpha, g) + \int_\alpha^\beta I_\beta(\alpha, f, \beta', g)d\beta', \\ &I(\alpha, f, \beta, g) \cong 0. \end{aligned} \tag{39}$$

The formulas (29), (30), (31), (32), (34), (35), (37), (38), and (39) prove that $I(\alpha, f, \beta, g)$, its first derivatives and its second derivatives of the type

$\partial^2/\partial\alpha\partial\beta$, $\partial^2/\partial\alpha\partial g$, $\partial^2/\partial\beta\partial f$, $\partial^2/\partial f\partial g$ may be expressed in terms of a, b, c, their first and second derivatives of the same type, and integrals over products of such functions.

Henceforth the definition (24) *of I is to be replaced by the explicit formula whose abbreviated equivalent is* (39), and which retains sense in the case that a, b, c admit only of continuous first derivatives and of continuous second derivatives of the indicated type. If we uniformly approximate a, b, c and said derivatives by polynomials in α, f, β, g and their respective derivatives, we may easily see that all of the formulas (29)–(32), (34), (35), (37)–(39) remain valid under the new assumptions and that $I(\alpha, f, \beta, g)$ still retains continuous first and second derivatives of said type.

Moreover a study of the dependence of I on the function $f(\alpha)$ shows that convergence of a sequence of continuously differentiable functions $f_\mu(\alpha)$ to a limit function $f_0(\alpha)$ implies the convergence of the corresponding functionals $I_\mu(\alpha, f, \beta, g)$ to a limit functional $I_0(\alpha, f, \beta, g)$, and uniform convergence of $f_\mu(\alpha)$ to $f_0(\alpha)$ entails uniform convergence of I_μ to I_0. In fact, $f_\mu(\alpha)$ appears in the definition of $I_\mu(\alpha, f, \beta, g)$ only in limits of integration with respect to f' or g', which implies the convergence mentioned of I_μ to I.

From the formulas (29)–(32), (34), (35), (37)–(39) we can derive estimates for $I(\alpha, f, \beta, g)$ and its derivatives. Suppose first that in $|\alpha|$, $|f|$, $|\beta|$, $|g| \leqq \omega$

$$|a|,\ |b|,\ |c| \leqq K,\quad |a_\alpha|, \cdots, |c_g| \leqq K',\quad |a_{\alpha\beta}|, \cdots, |c_{fg}| \leqq K''.$$

The terms suppressed in the above formulas by the use of the symbol $\cong$ are simple, double, triple, and quadruple integrals of polynomials of third degree in $a, b, c, a_\alpha, \cdots, c_{fg}$ with ranges of integration, respectively, $\leqq \rho\omega, \rho\omega^2, \rho\omega^3, \rho\omega^4$, where ρ denotes a sufficiently large number, for instance, 64. On the other hand, any one of the polynomials to be integrated is numerically smaller than a suitable polynomial in K, K', K'' of third degree with positive coefficients. Hence there exists a polynomial p with positive coefficients and of third degree in K, K', K'', such that for any one of formulas (29), $\cdots$, (39) the terms suppressed by the symbol $\cong$ are numerically less than or equal to

$$p(K, K', K'')(\omega + \omega^2 + \omega^3 + \omega^4).$$

We apply these estimates to the case where a, b, c *depend only indirectly* on α, f, β, g and are functions of $\psi_1, \cdots, \psi_n(\alpha, f, \beta, g)$ having third derivatives with respect to $\psi_1, \cdots, \psi_n$. We assume that in $|\alpha|$, $|f|$, $|\beta|$, $|g| \leqq \omega$, the following quantities exist and satisfy

$$|\psi_1|, \cdots, |\psi_n| \leqq k, \tag{40}$$

$$|\psi_{1\alpha}|, \cdots, |\psi_{ng}| \leqq k',\quad |\psi_{1\alpha\beta}|, \cdots, |\psi_{ngf}| \leqq k'',\quad (k, k', k'' > 0), \tag{41}$$

and that for ψ_i satisfying (40)

$$
\begin{aligned}
&|a|,\ |b|,\ |c| \leqq L,\\
&\left|\frac{\partial a}{\partial\psi_i}\right|,\ \left|\frac{\partial b}{\partial\psi_i}\right|,\ \left|\frac{\partial c}{\partial\psi_i}\right| \leqq L',\\
&\left|\frac{\partial^2 a}{\partial\psi_i\partial\psi_j}\right|, \cdots, \left|\frac{\partial^2 c}{\partial\psi_i\partial\psi_j}\right| \leqq L'',\\
&\left|\frac{\partial^3 a}{\partial\psi_i\partial\psi_j\partial\psi_l}\right|, \cdots, \left|\frac{\partial^3 c}{\partial\psi_i\partial\psi_j\partial\psi_l}\right| \leqq L'''.
\end{aligned}
\tag{42}
$$

We then write

$$I(\alpha, f, \beta, g) = I\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix}.$$

In order to utilize the estimates found, it is legitimate to replace K by L, K' by $nL'k'$, K'' by $nLk''+n^2L''k'^2$. Thus from (29), $\cdots$, (39) we obtain

$$
\begin{aligned}
&\left|I\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix}\right|,\ |I_\alpha|, \cdots, |I_g|\\
&\qquad\leqq (\omega+\omega^2+\omega^3+\omega^4)q(L, L', L'', k, k', k'')\\
&|I_{\alpha\beta}|,\ |I_{\alpha g}|,\ |I_{\beta f}|,\ |I_{fg}|\\
&\qquad\leqq n^2LL'^2k'^2+(\omega+\omega^2+\omega^3+\omega^4)q(L, \cdots, k''),
\end{aligned}
\tag{43}
$$

where q is a suitable polynomial with positive coefficients.

We next state bounds for the difference between

$$I\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} \quad\text{and}\quad I\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}$$

and its derivatives, assuming ψ_i' to be another system of continuously differentiable functions of α, f, β, g which satisfy the same inequalities (40), (41) as the ψ_i themselves do. For the sake of simplicity we suppose $\omega<\Omega$, with $\Omega>0$ and denote by u and v

$$u = \sum_{i=1}^n \max|\psi_i-\psi_i'| + \sum_{i=1}^n \max|\psi_{i\alpha}-\psi_{i\alpha}'| + \cdots + \sum_{i=1}^n \max|\psi_{ig}-\psi_{ig}'|,$$

$$
\begin{aligned}
v = {}&\sum_{i=1}^n \max|\psi_{i\alpha\beta}-\psi_{i\alpha\beta}'| + \sum_{i=1}^n \max|\psi_{i\alpha g}-\psi_{i\alpha g}'|\\
&+\sum_{i=1}^n \max|\psi_{i\beta f}-\psi_{i\beta f}'| + \sum_{i=1}^n \max|\psi_{ifg}-\psi_{ifg}'|,
\end{aligned}
$$

the maxima to be taken for the domain $|\alpha|, |f|, |\beta|, |g| \leqq \omega$.

We find by a procedure similar to the one previously used,

$$
(44.1)\qquad \left| I\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - I\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right| \leqq (u+v)C\omega,
$$

$$
\left| I_\alpha\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - I_\alpha\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right| \leqq (u+v)C\omega,
$$

. . . .

$$
\left| I_g\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - I_g\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right| \leqq (u+v)C\omega,
$$

$$
(44.2)\qquad \left| I_{\alpha\beta}\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - I_{\alpha\beta}\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right| \leqq uC + vC\omega,
$$

. . . .

$$
\left| I_{fg}\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - I_{fg}\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right| \leqq uC + vC\omega,
$$

with C depending on Ω, k, k', k'', L, L', L'', L'''. In an analogous way the difference between

$$
I\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} \quad\text{and}\quad \bar{I}\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}
$$

may be estimated under the assumption that the functions used in the formation of $\bar{I}$, namely $\bar{a}$, $\bar{b}$, $\bar{c}(\psi_1, \cdots, \psi_n)$, and their first and second derivatives, differ by less than ϵ from those used for

$$
I\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix}.
$$

We get

$$
\left| I\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - \bar{I}\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right| \leqq (u+v+\epsilon)C\omega,
$$

$$
\left| I_\alpha\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - \bar{I}_\alpha\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right|, \cdots, \leqq C(u+v+\epsilon)\omega,
$$

$$
(45)\qquad \left| I_{\alpha\beta}\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - \bar{I}_{\alpha\beta}\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right|, \cdots,
$$

$$
\left| I_{\beta g}\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - \bar{I}_{fg}\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right| \leqq C(\epsilon+u) + vC\omega.
$$

We now are able to formulate the main theorem of this section. Denote by

$$I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} \quad \text{the functional} \quad I\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix}$$

formed for

$$a = \psi_j, \qquad b = \psi_l, \qquad c = c_{ijl}.$$

THEOREM 4. *Denote by $\omega > 0$ a number $< \Omega$ and by D the domain $|\alpha|$, $|f|$, $|\beta|$, $|g| \leqq \omega$. Suppose c_{ijl} to be continuously differentiable up to the third order with respect to its arguments $\psi_1, \cdots, \psi_n$, and that for ψ_i in* (40), *$a = \psi_j$, $b = \psi_l$, $c = c_{ijl}$ the relations* (42) *hold. Suppose furthermore that in D the functions $\psi_i^0(\alpha, f, \beta, g)$ are continuously differentiable and admit of continuous second derivatives of the type mentioned such that*

$$|\psi_i^0(\alpha, f, \beta, g)| \leqq k/2, \tag{46}$$

$$|\psi^0{}_\alpha|, \cdots, |\psi^0{}_g| \leqq k'/2, \tag{47}$$

$$|\psi_{i\alpha\beta}^0|, |\psi_{i\alpha g}^0|, |\psi_{i\beta f}^0|, |\psi_{igf}^0| \leqq k''/3. \tag{48}$$

Then the system

$$\psi_i(\alpha, f, \beta, g) = \psi_i^0(\alpha, f, \beta, g) + \sum_{j,l=1}^{n} I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix}, \quad (i = 1, 2, \cdots, n), \tag{49}$$

has a solution $\psi_i(\alpha, f, \beta, g)$ existing and uniquely determined in $|\alpha|$, $|f|$, $|\beta|$, $|g| \leqq \omega'$, where $\omega' > 0$ is a number $\leqq \omega$ that may be determined with the aid of $\Omega, k, k', k'', L, L', L'', L'''$ only. $\psi_{i\alpha}(\alpha, f, \beta, g), \cdots, \psi_{ig}(\alpha, f, \beta, g)$ exist, are continuous and in absolute value $\leqq k'$, and $\psi_{i\alpha\beta}(\alpha, f, \beta, g), \cdots, \psi_{ifg}(\alpha, f, \beta, g)$ exist, are continuous and in absolute value $\leqq$*max $(k'', 12n^5LL'^2k'^2)$.

We start the proof by increasing, if necessary, k'' so as to satisfy the inequality

$$4n^5LL'^2k'^2 < k''/3.$$

We use successive approximations:

$$\psi_i^1(\alpha, f, \beta, g) = \psi_i^0(\alpha, f, \beta, g) + \sum_{k,l=1}^{n} I_{ikl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^0\end{pmatrix},$$

and generally for $m \geqq 0$

$$\psi_i^{m+1}(\alpha, f, \beta, g) = \psi_i^0(\alpha, f, \beta, g) + \sum_{k,l=1}^{n} I_{ikl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^m\end{pmatrix}. \tag{49.1}$$

* In particular ω' does not depend on the function $f(\alpha)$ used in the definition of I_{ijl}.

Denoting, as before, by C a constant depending only on Ω, k, k', k'', L, L', L'', by ∂ a generic first derivative with respect to α, f, β, or g, and by ∂^2 a generic derivative of type $\partial^2/\partial\alpha\partial\beta$, $\partial^2/\partial\alpha\partial g$, $\partial^2/\partial\beta\partial f$, $\partial^2/\partial f\partial g$, we set, for $m>0$,

$$u_m = \sum_{i,j,l=1}^{n} \max \left| I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^m\end{pmatrix} - I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^{m-1}\end{pmatrix}\right|$$

$$+ \sum_{i,j,l=1}^{n} \sum_{\partial} \max \left| \partial(I_{ikl}(\psi^m) - I_{ijl}(\psi^{m-1}))\right|$$

$$v_m = \sum_{i,j,l=1}^{n} \sum_{\partial^2} \max \left| \partial^2\left(I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^m\end{pmatrix} - I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^{m-1}\end{pmatrix}\right)\right|,$$

where the maxima are to be taken in $|\alpha|$, $|f|$, $|\beta|$, $|g| \leqq \omega'$ with $\omega' < \Omega$ to be determined later. On putting

$$u_0 = \sum_{i,j,l=1}^{n} \max \left| I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^0\end{pmatrix}\right| + \sum_{i,j,l=1}^{n} \sum_{\partial} \max \left| \partial I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^0\end{pmatrix}\right|,$$

$$v_0 = \sum_{i,j,l=1}^{n} \sum_{\partial^2} \max \left| \partial^2 I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^0\end{pmatrix}\right|,$$

we find, by (43),

$$u_0 \leqq C\omega', \tag{50}$$

$$v_0 \leqq 4n^5LL'^2k'^2 + C\omega', \tag{51}$$

and, for u_m and v_m the recursion formulas, in view of (44.1) and (44.2),

$$u_{m+1} \leqq C\omega'(u_m + v_m), \tag{52.1}$$

$$v_{m+1} \leqq Cu_m + C\omega' v_m, \qquad (m = 0, 1, 2, \cdots) \tag{52.2}$$

provided, however, that we can choose $\omega' > 0$ so as to make sure the existence of *all* successive approximations $|\alpha|$, $|f|$, $|\beta|$, $|g| \leqq \omega'$ in the common domain D' ($|\alpha|$, $|f|$, $|\beta|$, $|g| \leqq \omega'$). Now determine $\omega' < \Omega$ so small that

$$1 - C\omega' > 0, \qquad (1 - C\omega')^2 - C^2\omega' > 0,$$

$$U \equiv \frac{C\omega'(1 - C\omega') + (4n^5LL'^2k'^2 + C\omega')C\omega'}{(1 - C\omega')^2 - C^2\omega'} \leqq \frac{\min(k, k')}{2}, \tag{53}$$

$$V \equiv \frac{C^2\omega' + (1 - C\omega')(4n^5LL'^2k'^2 + C\omega')}{(1 - C\omega')^2 - C^2\omega'} \leqq \frac{2k''}{3}.$$

Note that

$$(1 + U + V)C\omega' \leqq U$$

and

$$4n^5LL'^2k'^2 + C\omega' + CU + C\omega'V \leqq V.$$

Now the conditions (46), (47), and (48) permit the construction of

$$I_{ikl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi^0\end{pmatrix}$$

in D, hence a fortiori in D' ($|\alpha|, |f|, |\beta|, |g| \leqq \omega'$), and we certainly have, by (53),

$$u_0 \leqq U,$$
$$v_0 \leqq V.$$

Suppose that we could construct, throughout D', the mth approximation $\psi_i^m(\alpha, f, \beta, g)$ and that

$$\sum_0^m u_j \leqq U, \tag{54}$$

$$\sum_0^m v_j \leqq V. \tag{55}$$

We are then able to prove that we can construct the $(m+1)$st approximation and that

$$\sum_0^{m+1} u_j \leqq U, \qquad \sum_0^{m+1} v_j \leqq V.$$

In fact, we have, in view of (54) and (55), (46), (47), and (48),

$$\begin{aligned} |\psi_i^{m+1}(\alpha, f, \beta, g)| &\leqq |\psi_i^0(\alpha, f, \beta, g)| + \sum_0^m u_j \\ &\leqq |\psi_i^0(\alpha, f, \beta, g)| + U \leqq k, \end{aligned} \tag{56}$$

$$|\partial\psi_i^{m+1}(\alpha, f, \beta, g)| \leqq |\partial\psi_i^0(\alpha, f, \beta, g)| + \sum_0^m u_j \leqq k', \tag{57}$$

$$|\partial^2\psi_i^{m+1}(\alpha, f, \beta, g)| \leqq |\partial^2\psi_i^0(\alpha, f, \beta, g)| + \sum_0^m v_j \leqq k'', \tag{58}$$

and by (52),

$$\sum_0^{m+1} u_j \leqq u_0 + C\omega'\left(\sum_0^m u_j + \sum_0^m v_j\right) \leqq C\omega'(1 + U + V) \leqq U,$$

$$\sum_0^{m+1} v_j \leqq v_0 + C\sum_0^m u_j + C\omega'\sum_0^m v_j \leqq 4n^5LL'^2k'^2 + C\omega' + CU + C\omega'V \leqq V.$$

Thus (54) and (55) hold for all $m \geqq 0$, and we conclude the uniform convergence of $\psi_i^m(\alpha, f, \beta, g)$, $\partial\psi_i^m$, $\partial^2\psi_i^m$ to limit functions $\psi_i(\alpha, f, \beta, g)$ and their corresponding derivatives $\partial\psi_i$, $\partial^2\psi_i$. The continuity relation (45) finally proves

$$I\begin{pmatrix}\alpha, f, \beta, g\\ \psi^m\end{pmatrix} \to I\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix}.$$

Hence by passage to the limit in (49.1) we obtain (49).

The uniqueness follows similarly from the relations analogous to (44.1) and (44.2),

$$u \leqq C\omega'(u + v) \tag{59.1}$$

$$v \leqq Cu + C\omega' v, \tag{59.2}$$

which yield $u(1 - C\omega') \leqq C\omega' v \leqq C\omega' \cdot Cu/(1 - C\omega')$, $u((1 - C\omega')^2 - C^2\omega') \leqq 0$, $u = 0$, $v = 0$, with

$$u = \sum_{i,j,l=1}^{n} \max \left| I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix} - I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi'\end{pmatrix}\right|$$
$$+ \sum_{i,j,l} \sum_{\partial} \max \left| \partial(I_{ijl}(\psi) - I_{ijl}(\psi'))\right|,$$
$$v = \sum_{i,j,l} \sum_{\partial^2} \max \left| \partial^2(I_{ijl}(\psi) - I_{ijl}(\psi'))\right|,$$

$\psi_i(\alpha, f, \beta, g)$ and $\psi_i'(\alpha, f, \beta, g)$ being solutions of (49).

Corollary 1. *If, in Theorem* 4, $f(\alpha)$, $\psi_1^0(\alpha, f, \beta, g), \cdots, \psi_n^0(\alpha, f, \beta, g)$, $\partial\psi_1^0(\alpha, f, \beta, g), \cdots, \partial\psi_n^0(\alpha, f, \beta, g)$, $\partial^2\psi_1^0(\alpha, f, \beta, g), \cdots, \partial^2\psi_n^0(\alpha, f, \beta, g)$, *and* $c_{ijl}(\psi_1, \cdots, \psi_n)$ *and its derivatives up to the third order depend on a parameter* μ *and converge uniformly as* $\mu \to \infty$, *then* $\psi_1(\alpha, f, \beta, g), \cdots, \psi_n(\alpha, f, \beta, g)$, $\partial\psi_1, \cdots, \partial\psi_n, \partial^2\psi_1, \cdots, \partial^2\psi_n$ *converge uniformly, as* $\mu \to \infty$, *in* $|\alpha|$, $|f|$, $|\beta|$, $|g| \leqq \omega'' \leqq \omega'$, *where* ω'' *depends only on* $\Omega, k, k', k'', L, L', L''$.

Denote by Δ the operation of taking the difference for two sufficiently large values of μ, and put, in the successive approximations of the proof of Theorem 4,

$$u_m = \sum_{i=1}^{n} \left| \Delta\psi_i^m(\alpha, f, \beta, g)\right| + \sum_{i=1,\partial}^{n} \left| \Delta\partial\psi_i^m\right|,$$

$$v_m = \sum_{i=1,\partial^2}^{n} \left| \Delta\partial^2\psi_i^m(\alpha, f, \beta, g)\right|.$$

Observing that $f(\alpha)$ enters in the functional I only as a limit of integration, as has been remarked earlier, we may use (45) and find, with some

$C=C(\Omega, k, k', k'', L, L', L'', L''')$ and a new and smaller value ω'' of ω', satisfying (53) with the new C:

$$u_m \leqq C(u_{m-1} + v_{m-1} + \epsilon)\omega'' + C\epsilon + u_0,$$
$$v_m \leqq Cu_{m-1} + Cv_{m-1}\omega'' + C\epsilon + v_0,$$
$$u_0 \leqq \epsilon, \qquad v_0 \leqq \epsilon.$$

Hence, for $m \to \infty$, $\lim u_m = u$, $\lim v_m = v$

$$u \leqq C(u + v + \epsilon)\omega'' + C\epsilon + u_0,$$
$$v \leqq Cu + Cv\omega'' + C\epsilon + v_0,$$
$$u(1 - C\omega'') \leqq (C + 1)\epsilon + \epsilon C\omega'' + C\omega'' v,$$
$$C\omega'' v \leqq \frac{C\omega''}{1 - C\omega''} Cu + \frac{(C+1)\epsilon C\omega''}{1 - C\omega''}\cdot$$

Hence u and v are $\leqq C'\epsilon$, with $C'(C, \omega'')$, which proves the corollary.

COROLLARY 2. *If, in Theorem* 4, $f(\alpha)$, $\psi_i^0(\alpha, f, \beta, g)$, C_{ijl} *depend on a parameter* μ, *and if* $f(\alpha)$ *converges uniformly as* $\mu \to \mu_0$, *then there exists a subsequence of* μ, *such that* $\psi_i(\alpha, f, \beta, g)$ *and also* $\psi_i(\alpha, f(\alpha), \beta, g(\beta))$ *converge uniformly.*

For by (57), $|\partial\psi_i(\alpha, f, \beta, g)| \leqq k'$ and hence the $\psi_i(\alpha, f, \beta, g)$ are equicontinuous and bounded, in view of (56). Hence there exists a uniformly convergent subsequence, and the corollary follows. From (33) we conclude under the hypothesis of the theorem, that

$$(60) \quad \frac{d^2}{d\alpha d\beta}\psi_i(\alpha, f(\alpha), \beta, g(\beta)) = \frac{d^2}{d\alpha d\beta}\psi_i^0(\alpha, f(\alpha), \beta, g(\beta)) + \sum_{j,l=1}^{n} c_{ijl}(\psi)\frac{d\psi_j}{d\alpha}\frac{d\psi_l}{d\beta},$$

and, for $\alpha=\beta$, $f=f(\alpha)$, $g=f(\alpha)$

$$\psi_i = \psi_i^0, \qquad \partial\psi_i = \partial\psi_i^0.$$

In the application we intend to make of Theorem 4 and its corollaries, the values of the constants such as Ω, k, k', k'', L, L', L'', L''' are of no importance. What matters, however, is their existence and their interdependence. Therefore, we are led to use the following terminology: we call a function *bounded* if its absolute value is bounded by a positive number irrespective of the values of its arguments and possible other parameters; we call, in a theorem, a quantity *relatively bounded* if its absolute value can be bounded by a positive number which depends only on other bounds *previously* introduced in the theorem; and we use the same term, in a proof, as meaning limitable by bounds, either assumed by the hypotheses of the theorem, or previously introduced in the course of the same proof.

Thus Theorem 4' and Corollary 2 may be formulated as follows:

THEOREM 5. *Suppose c_{ijl} and its derivatives up to the third order with respect to its arguments $\psi_1, \cdots, \psi_n$ bounded for bounded values of ψ_i, and assume that $\psi_i=\psi_i^0(\alpha, f, \beta, g), i=1, \cdots, n$, has derivatives $\partial\psi_i^0$, $\partial^2\psi_i^0$ which are continuous and bounded when α, f, β, g are bounded. Then the system*

$$\psi_i(\alpha, f, \beta, g) = \psi_i^0(\alpha, f, \beta, g) + \sum_{j,l=1}^{n} I_{ijl}\begin{pmatrix}\alpha, f, \beta, g\\ \psi\end{pmatrix}, \qquad (i = 1, 2, \cdots, n),$$

has a solution $\psi_i(\alpha, f, \beta, g)$, continuous together with $\partial\psi_i$ and $\partial^2\psi_i$. This solution exists, is uniquely determined and is relatively bounded together with the derivatives $\partial\psi_i$, $\partial^2\psi_i$ for relatively bounded α, f, β, g. If, in addition, $f(\alpha)$, $\psi_i^0(\alpha, f, \beta, g)$ and c_{ijl} and its derivatives up to the third order depend on a parameter μ, and if $f(\alpha)$ converges uniformly as $\mu\rightarrow\mu_0$, then there exists a subsequence of μ such that the corresponding functions $\psi_i(\alpha, f, \beta, g)$ and $\psi_i(\alpha, f(\alpha), \beta, g(\beta))$ converge uniformly for relatively bounded values of α, f, β, g.

4. **Hyperbolic systems.** The results of the preceding section may be used for a study of Cauchy's problem* for the system

$$(61)\qquad \begin{aligned} &\sum_{j=1}^{n} a_{ij}(\phi_1, \cdots, \phi_n)\frac{\partial\phi_j(\alpha, \beta)}{\partial\alpha} = 0, && i = 1, 2, \cdots, m < n,\\ &\sum_{j=1}^{n} a_{ij}(\phi_1, \cdots, \phi_n)\frac{\partial\phi_j(\alpha, \beta)}{\partial\beta} = 0, && i = m+1, \cdots, n, \end{aligned}$$

in which a_{ij} and its partial derivatives up to the fourth order as well as the reciprocal value of the determinant $|a_{ij}|$ are bounded for bounded values of $\phi_1, \cdots, \phi_n$. The initial line is a bounded neighborhood of the origin on the line $\alpha=\beta$, and on it the unknown functions $\phi_1(\alpha, \beta), \cdots, \phi_n(\alpha, \beta)$ assume relatively bounded values $\xi_1(\alpha), \cdots, \xi_n(\alpha)$ which are continuously differentiable.

In view of the applications we subject the $\xi_1(\alpha), \cdots, \xi_n(\alpha)$ to the following condition:

Condition ϑ. $\xi_1(\alpha), \cdots, \xi_n(\alpha)$ depend continuously on α, and there exists a transformation

$$\zeta_i(\alpha) = \sum_{j=1}^{n} \xi_j(\alpha)\gamma_{ij}, \qquad \xi_j(\alpha) = \sum_{i=1}^{n} \Gamma_{ji}\zeta_i(\alpha),$$

* A study of this Cauchy problem with a view to enlarging the class of admissible initial conditions was undertaken by Margaret Gurney in her dissertation, Brown University, 1935 (unpublished).

with constant γ_{ij} of determinant ± 1 such that the derivatives of $\zeta_2(\alpha)$, $\zeta_3(\alpha), \cdots, \zeta_n(\alpha)$ are bounded.

Then there exists a relatively bounded solution $\phi_1(\alpha, \beta), \cdots, \phi_n(\alpha, \beta)$ of (61) in a relatively bounded (α, β)-neighborhood of the origin, assuming the given initial values, continuous in α, β, and continuously differentiable with respect to α and β.

It should be noticed that the essential content of the above statement lies in the fact that the derivative of $\zeta_1(\alpha)$ has no influence on the determination of the domain of existence.

The idea of the proof is to construct instead of functions $\phi_i(\alpha, \beta)$ other functions ψ_i of four arguments α, f, β, g which reduce to the solution of the initial problem in question for $f=\zeta_1(\alpha)$, $g=\zeta_1(\beta)$. In order to conform with the terminology formerly introduced, we henceforth shall identify $\zeta_1(\alpha)$ with $f(\alpha)$.

We try to satisfy the following conditions for functions $\psi_i{}^0(\alpha, f, \beta, g)$:

$$\text{(i)} \qquad \frac{d^2\psi_i^0}{d\alpha d\beta}(\alpha, f(\alpha), \beta, g(\beta)) = 0,$$

$$\text{(ii)} \qquad \psi_i^0(\alpha, f(\alpha), \alpha, f(\alpha)) = \xi_i(\alpha),$$

$$\text{(iii)} \qquad \begin{aligned} &\sum_{j=1}^{n} a_{ij}(\xi_1(\alpha), \cdots, \xi_n(\alpha))\Big[\overset{0}{\psi}_{i\alpha}(\alpha, f(\alpha), \alpha, f(\alpha)) \\ &\qquad\qquad + \psi_{if}(\alpha, f(\alpha), \alpha, f(\alpha))\frac{df(\alpha)}{d\alpha}\Big] = 0, \quad i \leqq m, \\ &\sum_{j=1}^{n} a_{ij}(\xi_1(\alpha), \cdots, \xi_n(\alpha))\Big[\overset{0}{\psi}_{i\beta}(\alpha, f(\alpha), \alpha, f(\alpha)) \\ &\qquad\qquad + \psi_{ig}(\alpha, f(\alpha), \alpha, f(\alpha))\frac{df(\alpha)}{d\alpha}\Big] = 0, \quad i > m. \end{aligned}$$

We first introduce $\psi_i^0(\alpha, f, \alpha, f)$ by

$$\text{(62)} \qquad \psi_i^0(\alpha, f, \alpha, f) = \sum_{k=2}^{n} \Gamma_{ik}\zeta_k(\alpha) + \Gamma_{i1}f, \qquad i = 1, 2, \cdots, n,$$

which yields

$$\text{(63)} \qquad \overset{0}{\psi}_{if}(\alpha, f, \alpha, f) + \overset{0}{\psi}_{ig}(\alpha, f, \alpha, f) = \Gamma_{i1}.$$

To determine ψ_{if}^0 and ψ_{ig}^0 we set up the system

$$\text{(64)} \qquad \begin{aligned} &\sum_{k=1}^{n} a_{ik}(\psi^0(\alpha, f, \alpha, f))\overset{0}{\psi}_{kf}(\alpha, f, \alpha, f) = 0, \qquad i = 1, \cdots, m, \\ &\sum_{k=1}^{n} a_{ik}(\psi^0(\alpha, f, \alpha, f))\overset{0}{\psi}_{kg}(\alpha, f, \alpha, f) = 0, \qquad i = m+1, \cdots, n, \end{aligned}$$

which together with (63), in view of the boundedness of $|a_{ik}|^{-1}$, determines $\psi^0_{if}(\alpha, f, \alpha, f)$ and $\psi^0_{ig}(\alpha, f, \alpha, f)$ as analytic functions of $a_{ik}(\psi^0(\alpha, f, \alpha, f))$ and thus as relatively bounded functions with relatively bounded and continuous total derivatives with respect to α and f.

From (ii)

$$\psi^0_{i\alpha}(\alpha, f, \alpha, f) + \psi^0_{i\beta}(\alpha, f, \alpha, f) = \sum_{k=2}^{n} \Gamma_{ik} \frac{d\zeta_k(\alpha)}{d\alpha}, \qquad i = 1, 2, \cdots, n,$$

and by (64) and (iii)

$$\begin{aligned} &\sum_{k=1}^{n} a_{ik}(\xi_1(\alpha), \cdots, \xi_n(\alpha))\psi^0_{i\alpha}(\alpha, f(\alpha), \alpha, f(\alpha)) = 0, && i \leqq m; \\ &\sum_{k=1}^{n} a_{ik}\psi^0_{i\beta}(\alpha, f(\alpha), \alpha, f(\alpha)) = 0, && i > m, \end{aligned} \tag{65}$$

which determine $\psi^0_{i\alpha}(\alpha, f(\alpha), \alpha, f(\alpha))$ and $\psi^0_{i\beta}(\alpha, f(\alpha), \alpha, f(\alpha))$ as continuous and relatively bounded functions of α for relatively bounded α.

We now put

$$A_i(\alpha, f) = \int_0^f \psi^0_{ig}(\alpha, f', \alpha, f')df'.$$

Obviously, $A_i(\alpha, f)$ has continuous and relatively bounded derivatives with respect to f and α.

Finally put

$$\begin{aligned} \Psi_i^0(\alpha, f, \beta, g) = \psi_i^0(\alpha, f, \alpha, f) &+ A_i(\beta, g) - A_i(\alpha, f) \\ &- \int_\beta^\alpha [\psi^0_{i\beta}(\alpha', f(\alpha'), \alpha', f(\alpha')) - A_{i\alpha}(\alpha', f(\alpha'))]d\alpha'. \end{aligned} \tag{66}$$

The reader will easily verify that the function $\Psi_i^0(\alpha, f, \beta, g)$, as defined by (66), has the following properties:

$$\begin{aligned} \Psi_i^0(\alpha, f, \alpha, f) &= \psi_i^0(\alpha, f, \alpha, f), \\ \Psi^0_{ig}(\alpha, f, \alpha, f) &= \psi^0_{ig}(\alpha, f, \alpha, f), \\ \Psi^0_{if}(\alpha, f, \alpha, f) &= \frac{d}{df}\psi_i^0(\alpha, f, \alpha, f) - \frac{dA_i(\alpha, f)}{df} = \psi^0_{if}(\alpha, f, \alpha, f), \\ \Psi^0_{i\beta}(\alpha, f(\alpha), \alpha, f(\alpha)) &= \psi^0_{i\beta}(\alpha, f(\alpha), \alpha, f(\alpha)), \\ \Psi^0_{i\alpha}(\alpha, f(\alpha), \alpha, f(\alpha)) &= \psi^0_{i\alpha}(\alpha, f(\alpha), \alpha, f(\alpha)). \end{aligned}$$

Hence we are justified in considering $\Psi_i^0(\alpha, f, \beta, g)$ as an extension of those elements of the unknown function $\psi_i^0(\alpha, f, \beta, g)$ which were used in the construction of $\Psi_i^0(\alpha, f, \beta, g)$, and we write $\Psi_i^0(\alpha, f, \beta, g) = \psi_i^0(\alpha, f, \beta, g)$. We have, moreover, $\partial^2\psi_i^0 = 0$ so that (i) is true. From (62) we obtain (ii). Formulas (64) and (65) give (iii).

Evidently, $|\psi_i^0(\alpha, f, \beta, g)|$ is less than an arbitrary positive number ϵ if the bounds of the initial data $\xi_1(\alpha), \cdots, \xi_n(\alpha)$ are sufficiently small and if α, f, β, g are relatively bounded.

On differentiating the first m equations of (61) with respect to β, and the last $(n-m)$ equations with respect to α and solving with respect to the mixed derivatives, we obtain a system of the form

$$\frac{\partial^2\phi_i(\alpha, \beta)}{\partial\alpha\partial\beta} = \sum_{j,l} c_{ijl}(\phi_1, \cdots, \phi_n) \frac{\partial\phi_j(\alpha, \beta)}{\partial\alpha} \frac{\partial\phi_l(\alpha, \beta)}{\partial\beta}, \tag{67}$$

where the $c_{ijl}(\phi_1, \cdots, \phi_n)$ have bounded derivatives up to the third order for bounded $\phi_1, \cdots, \phi_n$. Replacing ϕ by ψ, we solve

$$\psi_i(\alpha, f, \beta, g) = \psi_i^0(\alpha, f, \beta, g) + \sum_{j,l=1}^{n} I_{ijl}\begin{pmatrix}\alpha, f, \beta, g \\ \psi\end{pmatrix}$$

with the aid of Theorem 5. By (60) and (i) we have

$$\frac{d^2\psi_i(\alpha, f(\alpha), \beta, g(\beta))}{d\alpha d\beta} = \sum_{j,l} c_{ijl}(\psi(\alpha, f(\alpha), \beta, g(\beta))) \frac{d\psi_j(\alpha, f(\alpha), \beta, g(\beta))}{d\alpha} \frac{d\psi_l(\alpha, f(\alpha), \beta, g(\beta))}{d\beta}.$$

In view of (ii), $\phi_i(\alpha, \beta) = \psi_i(\alpha, f(\alpha), \beta, g(\beta))$ assumes the given initial values, satisfies (61) on $\alpha = \beta$, and has continuous derivatives with respect to α and β and continuous mixed second derivatives with respect to α and β.

A conclusion, familiar in the theory of hyperbolic equations shows that equations (61) are satisfied identically in α and β.

Thus we have established the following theorem:

THEOREM 6. *If in (61) a_{ij} and its partial derivatives up to the fourth order and the reciprocal value of the determinant $|a_{ij}|$ are bounded for bounded values of $\phi_1, \phi_2, \cdots, \phi_n$, and if the initial values of $\phi_i(\alpha, \beta)$ on $\alpha = \beta$ are relatively bounded in a bounded neighborhood of $\alpha = 0$ $(= \beta)$ and satisfy condition ϑ, then Cauchy's problem has a solution existing for all relatively bounded α, β. This solution has continuous derivatives with respect to α and β. If the initial values and the a_{ij} depend on a parameter μ and converge uniformly as $\mu \to \mu_0$, then there exists a subsequence of μ, for which the corresponding solutions $\phi_i(\alpha, \beta)$ converge uniformly.*

COROLLARY 1. *The solution of Theorem 6 is unique.**

Let $|\alpha| \leqq A$, $|\beta| \leqq B$ be the common domain D of existence of two solutions of our initial problem. Denote by $u, v, w(\tau)$

$$u(\tau) = \sum_{i=1}^{n} \max |\Delta\phi_i|, \qquad v(\tau) = \sum_i \max \left| \Delta \frac{\partial\phi_i}{\partial\alpha} \right|,$$

$$w(\tau) = \sum_i \max \left| \Delta \frac{\partial\phi_i}{\partial\beta} \right|,$$

where the operator Δ indicates the difference of the expression following Δ for the two solutions, and the maximum is to be taken on that segment of the line $\tau = |\alpha - \beta|$ which is contained in D. By (67) we have for a suitable constant K

$$u(\tau) \leqq \int_0^\tau (v + w) |d\tau|, \qquad u(0) = 0,$$

$$v(\tau) \leqq K \int_0^\tau (u + v + w) |d\tau|, \qquad v(0) = 0,$$

$$w(\tau) \leqq K \int_0^\tau (u + v + w) |d\tau|, \qquad w(0) = 0.$$

Hence

$$u + v + w \leqq (2K + 1) \int_0^\tau (u + v + w) |d\tau|,$$

and by the well known iteration $u \equiv v \equiv w \equiv 0$.

By reasoning very similar to the preceding it may be shown that the dependence of the initial data on a parameter such that the initial data of ϕ_i and those of $\partial\phi_i/\partial\alpha$, $\partial\phi_i/\partial\beta$ satisfy a Lipschitz condition of exponent 1 in the parameter implies a Lipschitz condition of exponent 1 in the solution. Furthermore, passing to the limit from difference quotient to derivative with respect to the parameter we obtain the following corollary:

COROLLARY 2. *If, in Theorem 6, the initial data of ϕ_i and those of $\partial\phi_i/\partial\alpha$, $\partial\phi_i/\partial\beta$ are continuously differentiable with respect to a parameter, the solution and its first derivatives with respect to α and β are also continuously differentiable with respect to the parameter, continuity being understood with respect to the parameter and variables.*

---

* Cf. Hadamard, *Leçons sur le Problème de Cauchy*, Paris, 1932, pp. 488–501.

THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIF.

DECOMPOSITIONS AND DIMENSION OF CLOSED SETS IN R^n*

BY
ARTHUR N. MILGRAM

1. **Introduction.** It is our purpose to give a characterization of the dimension of closed sets immersed in R^n (euclidean space of n dimensions). This is done in terms of certain properties of the decompositions of these sets into a countable infinity of closed sets. The results are well known for finite decompositions of compact sets, but have never been shown to be so intrinsic a property of dimension as to remain valid under countable decompositions. This may be due to the fact that the proof seems to require much of the technique and many of the results of Alexandroff,† which are of very recent development. We may state our principal result as follows: *A closed subset F of R^n is of dimension r if and only if there exists an $\epsilon>0$, such that F may be decomposed into the sum of a countable infinity of closed sets $F_1, F_2, \cdots, F_s, \cdots$ of diameter less than ϵ, for which* dim $F_i \cdot F_j \leqq r-1$, $i \neq j$, *but for any such decomposition there exists a pair of integers m and n such that* dim $F_m \cdot F_n = r-1$.

This result follows quite readily when we have proved the following: If F is a closed subset of R^n, p a point of F, $F_1, F_2, \cdots, F_s, \cdots$ a decomposition of F into closed sets, z^{n-r-1} a cycle in $S(p, \epsilon)-F$, which does not bound in $S(p, \epsilon)-F$ but does bound in $S(p, \epsilon)-F_i$, $i=1, 2, \cdots$, then there exists a pair of integers m and n such that dim $F_m \cdot F_n \cdot S(p, \epsilon) \geqq r-1$. From this we obtain an interesting result which may be considered a generalization of a theorem due to Miss Mullikin.‡ We show that the sum of a countable number of closed sets, no one of which separates R^n, and the dimension of whose intersections taken pairwise does not exceed $n-3$, cannot separate R^n.

The author takes this opportunity to express his gratitude to Professor J. R. Kline, whose suggestions and unfailing encouragement made this paper possible.

2. **Notation.** The notation and definitions used in the sequel are widely employed. For example, $\delta(M)$ refers to the diameter of a point set M, $\rho(M_1, M_2)$ to the distance between the sets M_1 and M_2, $S(M, \epsilon)$ to the set of points x such that $\rho(M, x)<\epsilon$. We shall denote the boundary of a point set M by $B(M)$.

* Presented to the Society, March 27, 1937; received by the editors April 4, 1937.

† Especially his article *Dimensionstheorie*, Mathematische Annalen, vol. 106 (1932), pp. 161–238.

‡ A. M. Mullikin, *Certain theorems relating to plane connected point sets*, these Transactions, vol. 24 (1922), pp. 144–162.

The superscript attached to a symbol representing a chain will denote the dimension of the chain. The complex composed of the simplices of a chain C^i will be denoted by $|C^i|$. Cycles and chains occurring in this article will be assumed to have integral coefficients, although most of the results are just as valid if chains modulo m ($m \geqq 0$) or rational chains are used. The relation expressing the fact that the cycle z^r is (is not) the boundary of a chain in the domain D will be written $z^r \cong 0$ in D (z^r non-$\cong 0$ in D).

3. **Dimension of simplices.** Some years ago Sierpiński* proved that no continuum M can be expressed as the sum of a countable number of closed and proper subsets of M whose intersections are mutually vacuous. Remembering that the vacuous set is of dimension -1 and applying this theorem to the one-simplex, we have as an immediate result:

COROLLARY. *No one-simplex can be decomposed into the sum of a countable number of closed sets of diameter less than $\epsilon > 0$, where ϵ is less than the diameter of the simplex, and the dimension of the intersection of any pair of these closed sets is minus one.*

When the theorem is stated in this form, one is led to anticipate the more general statement:

THEOREM 1. *No r-simplex T^r can be decomposed into the sum of a countable number of closed sets of diameter less than $\epsilon > 0$, where ϵ is less than the diameter of T^r and the dimension of the intersection of any pair of these closed sets is at most $r-2$.*

Proof. Suppose the theorem false. Letting $F_1, F_2, \cdots, F_s, \cdots$ denote the closed sets referred to in the statement of the theorem, we have

$$T^r = F_1 + F_2 + \cdots + F_s + \cdots$$

and

$$\dim F_i \cdot F_j \leqq r - 2, \quad \delta(F_i) < \epsilon, \quad i \neq j \qquad (i, j = 1, 2, \cdots).$$

Denote the totality of intersections $F_i \cdot F_j$ by $P_1, P_2, \cdots$ and their sum by $P = \sum_{i=1}^{\infty} P_i$. P_i is closed.

Since the sum of a countable infinity of closed sets of dimension at most $r-2$ is of dimension at most $r-2$, it follows that P cannot fill any domain† in T^r.‡ But in the complement of each F_i, since F_i is of diameter less than

* W. Sierpiński, *Un théorème sur les continus*, Tôhoku Mathematical Journal, vol. 13 (1918), pp. 300–304.

† A domain is a connected open set. See P. Alexandroff and H. Hopf, *Topologie*, vol. 1, p. 51.

‡ P. Urysohn, *Mémoire sur les multiplicités Cantoriennes*, Fundamenta Mathematicae, vol. 8 (1927), pp. 337–341.

$\delta(T^r)$, there exists a domain. From this it follows readily that at least two of the sets $F_i - P$ $(i=1, 2, \cdots)$ are non-vacuous.

Let p be a point of $F_{i_1} - P$, and q a point of $F_{i_2} - P$, $i_1 \neq i_2$. We may assume p and q to be interior points of T^r. Now P_1 is a closed set, and $\dim P_1 \leqq r-2$. Consequently p can be joined to q by a polygonal line t_1 in the interior of $T^r - P_1$.* We enclose t_1 in a domain D_1 whose closure does not meet P_1.

Suppose we have constructed the domains $D_1, D_2, \cdots, D_{i-1}$, where

1. $\bar{D}_k \subset \bar{D}_{k-1}$,
2. $\bar{D}_k \cdot P_k = 0$, and
3. $D_k \supset p+q$ $(k=1, 2, \cdots, i-1)$.

In the construction of D_i we observe that $\dim P_i \cdot D_{i-1} \leqq r-2$. Hence p and q can be joined by a polygonal line t_i lying in $D_{i-1} - P_i \cdot D_{i-1}$, which we then enclose in a domain D_i whose closure is contained in D_{i-1} and does not meet P_i.

We thus obtain the sequence of continua

$$\text{(a)} \qquad \bar{D}_1, \bar{D}_2, \cdots, \bar{D}_i, \cdots,$$

where (a) satisfy relations 1, 2, and 3 above. $\Pi = \bar{D}_1 \cdot \bar{D}_2 \cdot \cdots \cdot \bar{D}_i \cdot \cdots$ is a continuum containing p and q, and, as is easily seen, containing no point of P. But the decomposition of Π into the closed sets $\Pi \cdot F_i$ $(i=1, 2, \cdots)$,

$$\Pi = \sum_{i=1}^{\infty} \Pi \cdot F_i,$$

affords a contradiction to Sierpiński's theorem, since at least two of these sets ($\Pi \cdot F_{i_1}$ and $\Pi \cdot F_{i_2}$) are non-vacuous, whereas the intersection of any pair is vacuous. This contradiction establishes the theorem.

By a slight modification in the method of constructing the domains D_i (that is, by constructing D_i as a chain of regions whose diameters are less than $1/i$), we could have been taken Π to be an arc.† We should thus have obtained an incidental proof of the following theorem:

THEOREM 2. *The complement in an n-dimensional simplex (or in R^n) of the sum of a countable infinity of closed sets of dimension at most $n-2$ is arcwise connected.*‡

Although, as observed in Theorem 1, an r-simplex cannot be decomposed into small closed sets with mutual intersections of dimension at most $r-2$,

---

* Cf. P. Urysohn, loc. cit., p. 307.

† For a discussion of this method of characterizing an arc, see R. L. Moore, *On the foundations of plane analysis situs*, these Transactions, vol. 17 (1916), pp. 133–139.

‡ This was first proved for the case $n=2$, i.e., for the plane, by J. R. Kline, *Concerning the complement of a countable infinity of point sets of a certain type*, Bulletin of the American Mathematical Society, vol. 23 (1916–1917), pp. 290–292.

it can always be decomposed into a countable number of closed sets, of arbitrarily small diameter, whose intersections taken pairwise are of dimension at most $r-1$ (for example, by a simplicial subdivision). We may therefore characterize the dimension of a simplex in the following way:

THEOREM. *A simplex T is of dimension r if and only if for any ϵ, where $\delta(T)>\epsilon>0$, T may be decomposed into the sum of a countable infinity of closed sets $F_1, F_2, \cdots, F_s, \cdots$ of diameter less than ϵ, for which* $\dim F_i \cdot F_j \leqq r-1$, *$i \neq j$, but for any such decomposition there exists a pair of integers m and n such that* $\dim F_m \cdot F_n = r-1$.

This, as well as Theorem 1, is a special case of more general considerations to be developed independently in a following section. Its chief interest lies in the simple proof based entirely on set-theoretic considerations. The same, or similar, methods do not seem adequate for a treatment of closed sets in R^n.

4. **Some preliminary lemmas and considerations.** After this simple characterization of the dimension of a simplex, it is quite natural to define the dimension of a closed set F in an analogous fashion. The definition is an inductive one, where the vacuous set is defined to be of dimension minus one.

DEFINITION. A closed set F is said to be of *dimension r* if there exists an $\epsilon>0$, such that F may be decomposed into the sum of a countable infinity of closed sets $F_1, F_2, \cdots, F_s, \cdots$ of diameter less than ϵ for which $\dim F_i \cdot F_j \leqq r-1$, $i \neq j$, but for any such decomposition there exists a pair of integers m and n such that $\dim F_m \cdot F_n = r-1$.

It is our aim in the sequel to show the complete equivalence between this definition of dimension and the Menger-Urysohn definition applied to closed sets. To accomplish this we prove several lemmas and theorems. The first of these, Lemma A_r, is based on the notion of ϵ-modification and simple r-dimensional obstruction introduced by Alexandroff in his article *Dimensionstheorie*, previously referred to, and certain methods and theorems proved there. For the sake of completeness we shall define ϵ-modification and simple r-dimensional obstruction and state the results used in this paper.

DEFINITION. Given a chain K and a positive number ϵ, the chain K' will be called an *ϵ-modification of K* if to each simplex x of K there corresponds a chain y in K' satisfying the following conditions:

a. If $x_i^h \rightarrow \sum c^j x_j^{h-1}$ then $y_i^h \rightarrow \sum c^j y_j^{h-1}$.
b. If $h=0$ (that is, if x_i^h is a vertex), then $x_i^0 = y_i^0$.
c. The sum $|x_i^h| + |y_i^h|$ is contained in a sphere of radius ϵ.
d. $K = \sum a_i x_i$ implies $K' = \sum a_i y_i$, where the y_i's are the chains corresponding to the simplices x_i.

Remark. The ϵ-modification will be called simple if the bounding relations are taken modulo m, $m \geqq 0$, and only integral coefficients appear. But if the y_i^h are chains with rational coefficients, then this is called an ϵ-modification modulo 0.

From the definition of ϵ-modification it is a short step to prove that if K' is an ϵ-modification of K, then for every simplicial transformation f, of K' into K, where each vertex of y_i^h goes into some vertex of x_i^h,

$$f(K') = K. \tag{1}$$

Definition. $F \subset R^n$, and x is a point of F. F is said to be a *simple r-dimensional obstruction in the neighborhood of x* if there exists an $\epsilon > 0$ so that for every δ sufficiently small $S(x, \delta) - F$ contains an $n-r-1$ dimensional cycle modulo 0, which does not bound in $S(x, \epsilon) - F$.

We can now state Alexandroff's theorem:

Theorem. *The set F is r-dimensional in the sense of Menger-Urysohn if and only if F is a simple r-dimensional obstruction in the neighborhood of at least one point, but forms no simple k-dimensional obstruction in the neighborhood of any point, if $k > r$.*

We turn now to the proof of several preliminary lemmas.

Lemma A_r. *Suppose*
1. *D a domain in R^n,*
2. *F a set closed relative to D,*
3. *the dimension* of F at most equal to r,*
4. *z^{n-r-2} a cycle in $D-F$,*
5. *K^{n-r-1} a chain bounded by z^{n-r-2} in D,*

are given. Under these conditions, if ϵ_1 is positive and less than $\frac{1}{2}\rho(K^{n-r-1}, B(D))$, there exists in $S(K^{n-r-1}, \epsilon_1) - F$ a chain C^{n-r-1} bounded by z^{n-r-2}, such that $C^{n-r-1} - K^{n-r-1} \simeq 0$ in D.

Proof. Denote the set $F \cdot \overline{S(K^{n-r-1}, \epsilon_1)}$ by F'. F' is a compact subset of R^n of dimension at most r (conditions 2 and 3). From the theorem of Alexandroff just quoted, F' cannot form an $(r+i)$-dimensional obstruction in the neighborhood of any point. Hence given an $\epsilon_i > 0$, we can find an $\epsilon_{i+1} > 0$, $\epsilon_{i+1} < \epsilon_i$, such that if $z^{n-r-i-1}$ is a cycle in $R^n - F'$, and $\delta(z^{n-r-i-1}) < \epsilon_{i+1}$, then there exists a chain C_1^{n-r-i} satisfying the conditions:

(a) $$C_1^{n-r-i} \rightarrow z^{n-r-i-1} \qquad (\text{in } R^n - F'),$$

(b) $$\delta(C_1^{n-r-i}) < \tfrac{1}{2}\epsilon_i, \qquad (i = 1, 2, \cdots, n-r-1).$$

* Here, and in the sequel, whenever the term dimension is used it is to be understood in the sense of Menger-Urysohn.

Now given ϵ_1, we find successively the numbers $\epsilon_2, \epsilon_3, \cdots, \epsilon_{n-r}$, subject to these conditions.

Let $|K^{n-r-1}|$ be subdivided into simplices of diameter less than ϵ_{n-r}. We may assume that none of the vertices of the subdivided complex lie on F, for by an arbitrarily small displacement, leaving z^{n-r-2} intact, the complex can be made to satisfy this condition. We denote the chain obtained from the subdivision of $|K^{n-r-1}|$ by $(K^{n-r-1})'$, wherein the orientation of the simplices will be that induced by their carriers in K^{n-r-1}, and their coefficients will be the same as those of their carriers.

Consider any one-simplex of $(K^{n-r-1})'$, say x_k^1. Its boundary $\dot{x}_k^1$ is a zero-cycle of $R_n - F'$ whose diameter is less than ϵ_{n-r}. We can find a chain y_k^1 in $R^n - F'$ bounded by $\dot{x}_k^1$ and such that $\delta(y_k^1) < \frac{1}{2}\epsilon_{n-r-1}$, from the restrictions on $\epsilon_2, \epsilon_3, \cdots, \epsilon_{n-r}$. The vertices of x_k^1 are a subset of the vertices of y_k^1.

Suppose now that the chains y_k^i have been constructed and ordered to the simplices x_k^i in such a way as to preserve incidence relations. Assume moreover that

1°. $\delta(y_k^i) < \frac{1}{2}\epsilon_{n-r-i}$,

2°. $y_k^i \subset R^n - F'$,

3°. every vertex of x_k^i is a vertex belonging to y_k^i.

If x_k^{i+1} is a simplex of $(K^{n-r-1})'$ and $\sum c^i x_j^i$ is its boundary, it follows from 1° and 3° that

$$\delta\left(\sum c^i y_j^i\right) < \epsilon_{n-r-i}.$$

Also, from 2° and the preservation of incidence under the ordering, $\sum c^i y_j^i$ is a cycle in $R^n - F'$. Hence from the conditions on the ϵ_s, $(s = 1, 2, \cdots, n-r)$, and from (a) and (b), there exists a chain

$$y_k^{i+1} \to \sum c^i y_j^i \quad \text{in} \quad R^n - F',$$

such that

$$\delta(y_k^{i+1}) < \tfrac{1}{2}\epsilon_{n-r-i-1}.$$

Each vertex of x_k^{i+1} is a vertex of y_k^{i+1} (from 3°). We are careful throughout the above process to choose the y_j^i's corresponding to simplices of $(z^{n-r-2})'$ as the simplices themselves. This may be done since $(z^{n-r-2})'$ is contained in $R^n - F'$. Continuing this process we arrive at an ϵ_1-modification of the chain $(K^{n-r-1})'$ which we may denote by C_*^{n-r-1}.

If

$$(K^{n-r-1})' = \sum a^i x_i^{n-r-1},$$

then

$$C_*^{n-r-1} = \sum a^i y_i^{n-r-1},$$

and from the construction, we have

$$C_*^{n-r-1} \to (z^{n-r-2})' \quad \text{in} \quad R^n - F.$$

Since $C_*^{n-r-1} \subset S(K^{n-r-1}, \epsilon_1)$, we can replace F' by F in the preceding relation. There exists a chain C_{**}^{n-r-1} such that

$$C_{**}^{n-r-1} \to z^{n-r-2} - (z^{n-r-2})' \quad \text{in} \quad |z^{n-r-2}|.$$

(C_{**}^{n-r-1} may be obtained by the so-called cylinder construction,† on z^{n-r-2}, in which the base is subdivided into isomorphism with $(z^{n-r-2})'$ and the vertical lines are collapsed into points.)

The chain

$$C^{n-r-1} = C_*^{n-r-1} + C_{**}^{n-r-1}$$

satisfies the statement of the lemma. C^{n-r-1} is bounded by z^{n-r-2}, and we must show that $C^{n-r-1} - K^{n-r-1} \cong 0$. We do this in two steps:

(c) $$C_*^{n-r-1} - (K^{n-r-1})' \cong 0 \quad \text{in} \quad S(K^{n-r-1}, \epsilon_1).$$

(d) $$C_{**}^{n-r-1} + (K^{n-r-1})' - K^{n-r-1} \cong 0 \quad \text{in} \quad |K^{n-r-1}|.$$

Perform a simplicial transformation of C_*^{n-r-1} into $(K^{n-r-1})'$ in such a way that the vertices of y_k^i are transformed into the vertices of x_k^i $(i = 0, 1, \cdots, n-r-1)$. Then

$$f(C_*^{n-r-1}) = (K^{n-r-1})'$$

follows from the discussion about relation (1). If x is a point of C_*^{n-r-1}, then $\rho(x, f(x)) < \epsilon_1$. Therefore the straight line $xf(x)$‡ lies in $S(K^{n-r-1}, \epsilon_1)$. Let G^{n-r-1} be a chain isomorphic to C_*^{n-r-1}. Denote the cylinder formed by the product of the interval $I(0 \leqq t \leqq 1)$ and $|G_*^{n-r-1}|$ by $|G^{n-r}|$, and subdivide and orient $|G^{n-r}|$ so that

$$\dot{G}^{n-r} = G^{n-r-1} \times 0 - G^{n-r-1} \times 1 - Z^{n-r-2} \times I,$$

where z^{n-r-2} is isomorphic with $(z^{n-r-2})'$. We now cnstruct a continuous transformation g such that every point x of $G^{n-r-1} \times 0$ corresponds to its image y under the isomorphism between $G^{n-r-1} \times 0$ and C_*^{n-r-1}, while $x \times 1$ goes into $f(y)$, and $x \times t$, where $0 < t < 1$, goes into the point dividing the line $yf(y)$ in the ratio $t:(1-t)$. Then

$$g(G^{n-r}) \subset S(K^{n-r-1}, \epsilon_1),$$

† For a discussion of this method see P. Alexandroff and H. Hopf, *Topologie*, vol. 1, pp. 196–198.

‡ Here $xf(x)$ means the straight line from x to $f(x)$.

and

$$g(G^{n-r})^{\cdot} = g(\dot{G}^{n-r}) = C_{*}^{n-r-1} - (K^{n-r-1})' + g(Z^{n-r-2} \times I).$$

But $g(Z^{n-r-2}\times I)=0$, since g sends the $(n-r-1)$-chain $Z^{n-r-2}\times I$ into the $(n-r-2)$-chain $(z^{n-r-2})'$. This establishes relation (c).

To establish (d), it is sufficient to note that the left-hand member is the boundary of the cylinder on K^{n-r-1} with base subdivided into $(K^{n-r-1})'$ and vertical lines degenerated into points. Adding (c) and (d), and using the definition of C^{n-r-1}, we have

$$C^{n-r-1} - K^{n-r-1} \cong 0 \quad \text{in} \quad S(K^{n-r-1}, \epsilon_1).$$

This completes the proof of the lemma.

LEMMA B. *Assume*

1. *K a complex,*
2. *F a closed subset of K,*
3. *z^r an r-cycle on K which does not bound in $K-F$,*
4. *K_i^{r+1} $(i=1, 2)$ two chains of K bounded by z^r,*
5. *K^{r+2} a chain bounded by the cycle $K_1^{r+1}-K_2^{r+1}$.*

Then the intersection I of F and K^{r+2} contains a continuum M joining K_1^{r+1} and K_2^{r+1}.†

Proof. Suppose I contains no continuum M joining K_1^{r+1} and K_2^{r+1}. Then there exists an $\epsilon<0$, such that whenever K^{r+2} is subdivided into simplices of diameter less than ϵ, and H^{r+2} is the collection of those simplices having a non-vacuous intersection with I, no component of H^{r+2} joins K_1^{r+1} and K_2^{r+1}. Assume the contrary. There exists a sequence of simplicial subdivisions of K^{r+2}, say $K_1^{r+2}, K_2^{r+2}, \cdots, K_i^{r+2}, \cdots$ where the simplices of K_i^{r+2} are of diameter less than $\epsilon_i>0$, $\lim_{i\to\infty}\epsilon_i=0$, and the collection H_i^{r+2}, of simplices of K_i^{r+2} meeting I, contains a component joining K_1^{r+1} and K_2^{r+1}. Denote these components by $L_1, L_2, \cdots, L_i, \cdots$. We can choose a subsequence of the L_i for which the limit inferior‡ is non-vacuous. Since $L_i \subset H_i^{r+2} \subset S(I, \epsilon_i)$, the limit superior of our subsequence is a subset of I. From a well known theorem of L. Zoretti,§ the limit superior is a continuum M. Moreover, since the L_i

† This lemma was stated and proved first by R. L. Wilder, *Generalized closed manifolds*, Annals of Mathematics, vol. 35 (1934), p. 879, for the case $K=R^n$ and $F=\Gamma^{n-r-1}$, a cycle linking z^r. In a footnote on the same page Wilder mentions the truth of the lemma when $K=R^n$ and F is any closed set. This latter statement would have been sufficient for our needs. However the formulation of Lemma C has intrinsic interest, and we offer it even though its full generality is not needed here.

‡ For a definition of the terms "limit inferior" and "limit superior" see Kuratowski, *Topologie* I, p. 152.

§ See, e.g., Hausdorff, *Mengenlehre*, 1927 edition, p. 163.

join K_1^{r+1} and K_2^{r+1}, M does likewise, contrary to our assumption of the falsity of the lemma.

Now let ϵ and H^{r+2} be defined as in the preceding paragraph. Let $'C^{r+2}$ denote the sum of the components of H^{r+2} which meet K_1^{r+1}. We form the following chain from the simplices of C^{r+2}: If x^{r+2} is a simplex of $'C^{r+2}$, then x^{r+2} is assumed to have the same orientation and coefficient as its carrier in K^{r+2}. We denote this chain by $(C^{r+2})'$. Similarly, form the chains $(K^{r+2})'$, $(K_1^{r+1})'$, $(K_2^{r+1})'$, the subdivisions of the chains K^{r+2}, K_1^{r+1}, K_2^{r+1}. If L^{r+1} is the boundary of $(C^{r+2})'$, we have

$$(C^{r+2})' \rightarrow L^{r+1}$$

and

$$(K^{r+2})' \rightarrow (K_1^{r+1})' - (K_2^{r+1})'.$$

L^{r+1} contains points of I only in those simplices which are contained in $(K_1^{r+1})'$. For $(C^{r+2})'$ does not meet $(K_2^{r+1})'$, and if x^{r+1} is a simplex of L^{r+1} containing a point of I, then all simplices having x^{r+1} on their boundary belong to $(C^{r+2})'$. Hence x^{r+1} can only occur with coefficient different from zero when it belongs to either $(K_1^{r+1})'$ or $(K_2^{r+1})'$. Those simplices of L^{r+1} which meet I belong to $(K_1^{r+1})'$ and have the same coefficient as they have in $(K_1^{r+1})'$. It follows that $(K_1^{r+1})'-L^{r+1}$ is a chain in $|K^{r+2}|-I$. But

$$L^{r+1} \rightarrow 0$$

and

$$(K_1^{r+1})' \rightarrow (z^r)'.$$

Hence

$$(K_1^{r+1})' - L^{r+1} \rightarrow (z^r)'$$

in $|K^{r+2}|-I$, and therefore in $K-F$. Since $z^r \cong (z^r)'$ in $|z^r|$, we have z^r does bound in $K-F$ contrary to the hypothesis of the lemma. This final contradiction completes the proof.

Let F be a subset of the domain $D \subset R^n$. F is closed relative to D and decomposed into the sets $F_1, F_2, \cdots, F_i, \cdots$, each closed relative to D. Denote the totality of products $F_i \cdot F_j$, $i \neq j$, by $P_1, P_2, \cdots, P_i, \cdots$, and their sum by $P=\sum_{i=1}^{\infty} P_i$. Let $H_i=F_i-P$ $(i=1, 2, \cdots)$.

LEMMA C. *If infinitely many of the sets H_i are non-vacuous, then infinitely many of them contain points in the complement of S, the limit superior of $H_1, H_2, \cdots, H_i, \cdots$.*†

† It would be sufficient for our purposes to show that at least two of the sets H' contain points in the complement of S, but the proof is the same in either case.

Proof. Suppose there were a number s such that $S \supset \sum_{i=s}^{\infty} H_i$. Let p_1 be a point of the first non-vacuous set, say H_{s_1}.

$$F_i \cdot p_1 = 0 \qquad (i = 1, 2, \cdots, s_1 - 1).$$

For if $F_{i'} \supset p_1$, $i' < s_1$, then $F_{i'} \cdot F_{s_1} = P_k \supset p_1$, and $H_{s_1} \cdot p_1 = p_1 \cdot (F_{s_1} - P) \subset p_1 \cdot (F_{s_1} - P_k) = 0$. Consequently $F - (F_1 + F_2 + \cdots + F_{s_1 - 1}) \supset p_1$. As $F_1 + F_2 + \cdots + F_{s_1-1}$ is closed relative to D, there is a neighborhood N_1 of p_1 such that $N_1 \subset D - (F_1 + F_2 + \cdots + F_{s_1-1})$.

Since $S \supset p_1$, N_1 contains a point $p_2 \subset H_{s_2}$, where $s_2 > s_1$. In the same way we find a neighborhood $N_2 \subset N_1$ such that $\overline{N}_2 \cdot (F_1 + F_2 + \cdots + F_{s_2-1}) = 0$. Continuing this process, we obtain a sequence of integers $s_1 < s_2 < \cdots, s_i \geqq i$, and neighborhoods $N_1, N_2, \cdots, N_i, \cdots$ such that

$$\overline{N}_{i+1} \subset N_i,$$

and

$$\overline{N}_i \cdot (F_1 + F_2 + \cdots + F_{s_i - 1}) = 0.$$

Since $\overline{N}_i$ is compact, $\prod_{i=1}^{\infty} \overline{N}_i \neq 0$. If x is a point of $\prod_{i=1}^{\infty} \overline{N}_i$, there is an integer t such that $F_{t-1} \supset x$. But x is also contained in $\overline{N}_t$, while $\overline{N}_t \cdot (F_1 + F_2 + \cdots + F_{s_t - 1}) = 0$. This contradiction establishes the lemma.

5. **Principal theorems.** Let F be a closed subset of R^n, p a point of F, and ϵ a positive number. $F_1, F_2, \cdots, F_s, \cdots$ is a decomposition of F into closed sets.

THEOREM 3. *If there exists in $S(p, \epsilon) - F$ a cycle z^{n-r-1}, which does not bound in $S(p, \epsilon) - F$ but does bound in $S(p, \epsilon) - F_i$, $i = 1, 2, \cdots$, then there is a pair of integers m and n, $m \neq n$, such that,*

$$\dim F_m \cdot F_n \cdot S(p, \epsilon) \geqq r - 1.$$

Proof. If we assume the existence of the cycle z^{n-r-1}, F contains a closed subset A which is irreducible with respect to the property

$$z^{n-r-1} \text{ non-} \cong 0 \quad \text{in} \quad S(p, \epsilon) - A. \tag{2}$$

This may be seen by first showing that when $F \supset F' \supset F'' \supset \cdots \supset F^{(k)} \supset \cdots$ are closed sets and z^{n-r-1}non-$\cong 0$ in $S(p, \epsilon) - F^{(k)}$, then $F_\omega = \prod_{k=1}^{\infty} F^{(k)}$ is likewise closed and z^{n-r-1}non-$\cong 0$ in $S(p, \epsilon) - F_\omega$. Then, from a well known theorem due to Brouwer,† the existence of A follows. To prove the first point it is sufficient to remark that if K^{n-r} is any chain in $S(p, \epsilon)$ bounded by z^{n-r-1}, then K^{n-r} has a compact and non-vacuous intersection with each of the sets $F^{(k)}$. The product of these intersections is non-vacuous and belongs to F_ω.

† See K. Menger, *Dimensionstheorie*, Leipzig and Berlin, 1928, p. 69.

Let $A_i = F_i \cdot A$. A_i is closed, and

$$A = \sum_{i=1}^{\infty} A_i. \tag{3}$$

From $F_i \supset A_i$ follows $S(p, \epsilon) - A_i \supset S(p, \epsilon) - F_i$. By hypothesis, z^{n-r-1} bounds in $S(p, \epsilon) - F_i$. Hence

$$z^{n-r-1} \cong 0 \quad \text{in} \quad S(p, \epsilon) - A_i \qquad (i = 1, 2, \cdots). \tag{4}$$

Also $A_i \cdot A_j \cdot S(p, \epsilon) \subset F_i \cdot F_j \cdot S(p, \epsilon)$, which gives

$$\dim A_i \cdot A_j \cdot S(p, \epsilon) \leqq \dim F_i \cdot F_j \cdot S(p, \epsilon).$$

From this point to the completion of the proof, we shall suppose the theorem false, that is, $\dim F_i \cdot F_j \cdot S(p, \epsilon) \leqq r-2$, for all i, j, $(i \neq j)$. Consequently

$$\dim A_i \cdot A_j \cdot S(p, \epsilon) \leqq r - 2, \quad \text{for} \quad i \neq j.$$

Arranging the countable collection of products $A_i \cdot A_j \cdot S(p, \epsilon)$ in a sequence, and renaming them $P_1, P_2, \cdots$, we have $P = \sum_{i=1}^{\infty} P_i$ is of dimension at most $r-2$. This is a consequence of the well known result that the sum of a countable number of sets, closed relative to a domain and each of dimension at most $r-2$, is itself of dimension at most $r-2$. We have thus constructed a set A and a subdivision $A_1, A_2, \cdots$ which bear the same relation to z^{n-r-1} and $S(p, \epsilon)$ as do F and its subdivision and in addition is irreducible with respect to the non-bounding of z^{n-r-1} in $S(p, \epsilon) - A$.

We now prove that if

$$A(t) = \sum_{i=1}^{t} A_i,$$

then

$$z^{n-r-1} \cong 0 \quad \text{in} \quad S(p, \epsilon) - A(t), \qquad t = 1, 2, \cdots. \tag{5}$$

This is demonstrated for the case $t = 2$. A simple induction then carries the demonstration to any finite t. For, if by assumption, z^{n-r-1} bounds in $S(p, \epsilon) - A(t-1)$, then $A(t-1)$ and A_t satisfy the same conditions as do A_1 and A_2, and $A(t-1) + A_t = A(t)$. ($A(t-1)$ is closed since it is the sum of a finite number of closed sets. Moreover, $A(t-1) \cdot A_t \cdot S(p, \epsilon) = A_1 \cdot A_t \cdot S(p, \epsilon) + A_2 \cdot A_t \cdot S(p, \epsilon) + \cdots + A_{t-1} \cdot A_t \cdot S(p, \epsilon)$, being the sum of a finite number of sets closed relative to $S(p, \epsilon)$, and each of dimension at most $r-2$, satisfies the relation $\dim A(t-1) \cdot A_t \cdot S(p, \epsilon) \leqq r-2$.) We turn to the proof for $t = 2$.

We can find chains C_j^{n-r} in $S(p, \epsilon) - A_j$, $j = 1, 2$, (relation (4)) such that

$$C_j^{n-r} \to z^{n-r-1} \quad \text{in} \quad S(p, \epsilon) - A_j, \qquad j = 1, 2.$$

Denote the cycle $C_1^{n-r}-C_2^{n-r}$ by Z^{n-r}. Let K^{n-r+1} be a chain bounded by Z^{n-r} in $S(p, \epsilon)$. $A_1 \cdot A_2 \cdot S(p, \epsilon)$, Z^{n-r}, K^{n-r+1}, $S(p, \epsilon)$ satisfy the same conditions as do F, Z^{n-r}, K^{n-r+1}, D, respectively, in Lemma A_{r-2}. Hence there exists a chain C^{n-r+1} such that

$$C^{n-r+1} \rightarrow z^{n-r} \quad \text{in} \quad S(p, \epsilon) - A_1 \cdot A_2.$$

Now if z^{n-r-1} did not bound in $S(p, \epsilon)-A(2)$, then C^{n-r+1}, $A(2)$, z^{n-r-1}, C_i^{n-r}, C^{n-r+1} would satisfy the same conditions as K, F, z^r, K_i^{r+1}, K^{r+2}, respectively, in Lemma B. Hence $A(2) \cdot |C^{n-r+1}|$ would contain a continuum M joining C_1^{n-r} and C_2^{n-r}. This is impossible.

$$M = M \cdot A_1 + M \cdot A_2.$$

Neither of these sets is vacuous, since $A_2 \supset M \cdot |C_1^{n-r}| \neq 0$ and $A_1 \supset M \cdot |C_2^{n-r}| \neq 0$. The intersection $(M \cdot A_1) \cdot (M \cdot A_2) = M \cdot (A_1 \cdot A_2) \subset |C^{n-r+1}| \cdot (A_1 \cdot A_2) = 0$. This negates the connectedness of M and therefore establishes relation (5).

Infinitely many of the sets $S(p, \epsilon) \cdot A_i - P$ are non-vacuous. Suppose to the contrary that only

$$S(p, \epsilon) \cdot A_1 - P, \ S(p, \epsilon) \cdot A_2 - P, \ \cdots, \ S(p, \epsilon) \cdot A_t - P$$

are non-vacuous. Then

$$S(p, \epsilon) \cdot A = S(p, \epsilon) \cdot A(t) + P.$$

But there exists a chain C^{n-r} (relation (5)) such that

$$C^{n-r} \rightarrow z^{n-r-1} \quad \text{in} \quad S(p, \epsilon) - A(t).$$

We then find a neighborhood D of C^{n-r} which is small enough to lie in $S(p, \epsilon)$ and to exclude $A(t)$. Since $A \cdot D = P \cdot D$, it would follow that dim $A \cdot D \leqq r-2$. Again applying Lemma A_{i-2}, z^{n-r-1} would bound in $D - A \cdot D$ and consequently in $S(p, \epsilon) - A$, contradicting relation (2).

Now $S(p, \epsilon) \cdot A = \sum_{i=1}^{\infty} S(p, \epsilon) \cdot A_i$, and infinitely many of the sets $S(p, \epsilon) \cdot A_i - P$ are non-vacuous. Hence, by Lemma C, there exist two points p_1 and p_2 belonging to different sets, say, $A_{i_1} - P$ and $A_{i_2} - P$, and such that neither p_1 nor p_2 belongs to the limit superior of $S(p, \epsilon) \cdot A_1 - P$, $S(p, \epsilon) \cdot A_2 - P, \cdots$. We can find a number α sufficiently small so that $\overline{S(p_k, \alpha)}$ will meet no set of $S(p, \epsilon) \cdot A_1 - P$, $S(p, \epsilon) \cdot A_2 - P, \cdots$ other than $A_{i_k} - P$, $(k=1, 2)$. The sets $A - S(p_k, \alpha)$ $(k=1, 2)$ are closed and proper subsets of A. It follows from the irreducibility of A that there exist chains C_k^{n-r} such that

$$C_k^{n-r} \rightarrow z^{n-r-1} \quad \text{in} \quad S(p, \epsilon) - \{A - S(p_k, \alpha)\}.$$

Choose a number ϵ_1 such that ϵ_1 is smaller than either of the numbers

$\frac{1}{2}\rho(C_k^{n-r}, A-S(p_k, \alpha))$. We now replace the chains C_k^{n-r} by chains C_{1k}^{n-r}, lying in $S(C_k^{n-r}, \epsilon_1)-P_1$, which are bounded by z^{n-r-1}. This is of course possible by Lemma A_{i-1}. (P_1 is closed relative to $S(C_k^{n-r}, \epsilon_1)$ and of dimension less than $r-1$.) The cycle $C_{11}^{n-r}-C_{12}^{n-r}$ lies in the complement of P_1 in $S(p, \epsilon)$ and hence bounds in $S(p, \epsilon)-P_1$. This may be shown by allowing $C_{11}^{n-r}-C_{12}^{n-r}$ to bound some chain in $S(p, \epsilon)$, and then displacing this chain to another chain C_1^{n-r+1} bounded by $C_{11}^{n-r}-C_{12}^{n-r}$ in $S(p, \epsilon)-P_1$; a permissible operation as shown by Lemma A_{r-2}.

Let us suppose that we have constructed the following chains:

$$C_1^{n-r+1},\ C_2^{n-r+1},\ \cdots,\ C_t^{n-r+1}. \tag{6}$$

Assume further that these chains satisfy the following conditions:

(a) $\rho(C_i^{n-r+1}, P_i) > \alpha_{ii} > 0$.

(b) $\alpha_{si} = \rho(C_s^{n-r+1}, P_i) > \frac{1}{2}\alpha_{ii}$ for $s > i$.

(c) $C_i^{n-r+1} \to C_{i1}^{n-r} - C_{i2}^{n-r}$ in $S(p, \epsilon)$.

(d) $C_{ik}^{n-r} \to z^{n-r-1}$ in $S(p, \epsilon) - \{A - S(p_k, \alpha)\}$, $k = 1, 2$.

(e) $d_i = \rho(C_i^{n-r+1}, B(S(p, \epsilon))) > d > 0$.

We proceed with the construction of C_{t+1}^{n-r+1}.

Choose a number δ smaller than the minimum of the numbers

a⁰. $\alpha_{ti} - \frac{1}{2}\alpha_{ii}$, $i = 1, 2, \cdots, t$,

b⁰. $\rho(C_{tk}^{n-r}, A - S(p_k, \alpha))$, $k = 1, 2$,

c⁰. $d_t - d$.

If $C_{t+1,k}^{n-r}$ is any chain bounded by z^{n-r-1} in $S(C_{tk}^{n-r}, \delta)$, then $C_{t+1,k}^{n-r}$ satisfies condition (d). To prove this we show that if x is a point of $A-S(p_k, \alpha)$ and y a point of $C_{t+1,k}^{n-r}$, then $\rho(x, y)>0$. $C_{t+1,k}^{n-r}$ is assumed to lie in $S(C_{t,k}^{n-r}, \delta)$, so that there is a point z of $C_{t,k}^{n-r}$ such that $\rho(y, z)<\delta$. Condition b⁰ applied to

$$\rho(x, y) \geqq \rho(x, z) - \rho(y, z)$$

yields

$$\rho(x, y) \geqq \rho(C_{t,k}^{n-r}, A - S(p_k,\alpha)) - \delta > 0.$$

If C_{t+1}^{n-r+1} is a chain in $S(C_t^{n-r+1}, \delta)$, C_{t+1}^{n-r+1} will satisfy (b), that is,

$$\rho(C_{t+1}^{n-r+1}, P_i) > \tfrac{1}{2}\alpha_{ii}, \qquad i < t + 1.$$

For if x is a point of C_{t+1}^{n-r+1}, y a point of P_i, there is a point z of C_t^{n-r+1} such that $\rho(x, z)<\delta$. From $\rho(x, y)\geqq\rho(z, y)-\rho(x, z)$ and condition a⁰ on the number δ, we have

$$\rho(x, y) \geqq \rho(C_t^{n-r+1}, P_i) - \delta > \alpha_{ti} - (\alpha_{ti} - \tfrac{1}{2}\alpha_{ii}) = \tfrac{1}{2}\alpha_{ii}.$$

Similar considerations show that C_{t+1}^{n-r+1} would satisfy (e).

In $S(C_{t,k}^{n-r}, \delta)-P_{t+1}$, $(k=1, 2)$, we can find a chain $C_{t+1,k}^{n-r}$, and in $S(C_{t,k}^{n-r}, \delta)$ a chain ${}^kC^{n-r-1}$ such that

$$C_{t+1,k}^{n-r} \to z^{n-r-1} \quad \text{in} \quad S(C_{t,k}^{n-r}, \delta) - P_{t+1},$$

$$(7) \qquad {}^kC^{n-r+1} \to C_{t+1,k}^{n-r} - C_{t,k}^{n-r} \quad \text{in} \quad S(C_{t,k}^{n-r}, \delta).$$

(This is justified by Lemma A_{r-2}.) The chain

$$(8) \qquad {}^*C_{t+1}^{n-r+1} = C_t^{n-r+1} + {}^1C^{n-r+1} - {}^2C^{n-r+1}$$

is such that

$${}^*C_{t+1}^{n-r+1} \to C_{t+1,1}^{n-r} - C_{t+1,2}^{n-r} \quad \text{in} \quad S(C_t^{n-r+1}, \delta)$$

by (7) and (c) applied to (8). Since the boundary of ${}^*C_{t+1}^{n-r+1}$ lies in $S(C_t^{n-r+1}, \delta)-P_{t+1}$, by Lemma A_{r-2} we can replace ${}^*C_{t+1}^{n-r+1}$ by C_{t+1}^{n-r+1}, a chain in $S(C_t^{n-r+1}, \delta)-P_{t+1}$ bounded by $C_{t+1,1}^{n-r}-C_{t+1,2}^{n-r}$.

The set P_{t+1} is closed relative to $S(p, \epsilon)$, and C_{t+1}^{n-r+1} is contained in $S(p, \epsilon)$ and does not meet P_{t+1}. Consequently we can find a number $\alpha_{t+1,t+1}$ for which the relation

$$\rho(C_{t+1}^{n-r+1}, P_{t+1}) > \alpha_{t+1,t+1} > 0$$

holds. We have thus obtained an extension of the system (6) by the addition of C_{t+1}^{n-r+1}. Since we had previously shown that (6) existed for $t=1$, this latter shows that (6) can be extended indefinitely. We may therefore suppose that we have constructed a countable infinity

$$(6') \qquad C_1^{n-r+1}, C_2^{n-r+1}, \cdots, C_t^{n-r+1}, \cdots$$

of chains satisfying relations (a)–(e) inclusive.

Consider any chain of the above sequence. $A \cdot |C_t^{n-r+1}|$ is a closed set, and z^{n-r-1} does not, of course, bound on $|C_t^{n-r+1}| - A \cdot |C^{n-r+1}|$. Applying Lemma B, we see that $A \cdot |C_t^{n-r+1}|$ contains a continuum M_t joining $C_{t,1}^{n-r}$ and $C_{t,2}^{n+r}$. Since these latter chains meet A only in $S(p_2, \alpha)$ and $S(p_1, \alpha)$ respectively, we have

$$M_t \cdot S(p_k, \alpha) \neq 0, \qquad k = 1, 2.$$

From $|C_t^{n-r+1}| \subset S(p, \epsilon - d)$ (condition (e)), it follows that

$$M_t \subset S(p, \epsilon - d),$$

that is, the sequence

$$(9) \qquad M_1, M_2, \cdots, M_t, \cdots$$

is uniformly bounded, and the limit superior of (9) is contained in the interior of $S(p, \epsilon)$. From (9) we choose a subsequence

$$M_{t_1}, M_{t_2}, \cdots, M_{t_s}, \cdots \tag{10}$$

of which the limit inferior is non-vacuous. From the theorem of L. Zoretti previously referred to, the limit superior of (10) is a continuum M. Since each M_i is contained in $\overline{S(p, \epsilon-d)}$, M is likewise. From $A \supset M_i$, and A closed, we have

$$A \supset M. \tag{11}$$

We affirm

$$M \cdot P = 0. \tag{12}$$

If we assumed the contrary, there would be some P_i for which $M \cdot P_i \neq 0$. But $M_s \subset |C_s^{n-r+1}|$, and from condition (b) on (6′) we should have

$$\rho(M_s, P_i) > \tfrac{1}{2}\alpha_{ii} > 0, \quad \text{for all} \quad s > i.$$

This leads to the conclusions

$$\rho(M, P_i) \geqq \tfrac{1}{2}\alpha_{ii},$$

and therefore $M \cdot P_i = 0$, which proves (12). Combining (3) and (11), we have

$$M = \sum_{i=1}^{\infty} M \cdot A_i.$$

Moreover

$$(M \cdot A_i) \cdot (M \cdot A_j) \subset M \cdot P = 0, \qquad i \neq j.$$

$M \cdot A_s$ is closed.

We have thus obtained a decomposition of the continuum M into a countable infinity of closed sets whose intersections taken pairwise are vacuous. This is impossible according to Sierpiński's theorem unless all but one of these sets are vacuous. But M must contain a point of $A \cdot \overline{S(p_k, \alpha)}$ $(k=1, 2)$, since each of the sets in (9) does. From the choice of the number α, we have $A_{i_k} - P \supset A \cdot \overline{S(p_k, \alpha)} - P$. M must therefore have a non-vacuous intersection with both A_{i_1} and A_{i_2}. This final contradiction completes the proof of the theorem.

Corollary. *In R^n let F' be a closed set and z_1^{n-r-1} a cycle in $R^n - F'$ which does not bound in $R^n - F'$. Moreover let $F' = F_1' + F_2' + \cdots + F_s' + \cdots$ where F_s' is closed $(s=1, 2, \cdots)$ and z_1^{n-r-1} bounds in $R^n - F_s'$. Then there exists a pair of integers m and n such that* dim $F_m' \cdot F_n' \geqq r-1$.

To show this we let p be any point of F' and ϵ some positive number. R^n is homeomorphic to $S(p, \epsilon)$, and the homeomorphism f may be so taken as to leave p invariant. If we take $F=\overline{f(F')}$, $F_s=\overline{f(F_s')}$ and $z^{n-r-1}=f(z_1^{n-r-1})$, then, since

$$F_i \cdot F_j \cdot S(p, \epsilon) = f(F_i') \cdot f(F_j'),$$

we have from Theorem 3,

$$\dim f(F_m') \cdot f(F_n') \geqq r - 1$$

for some pair of integers m and n. But because of the invariance of dimension under homeomorphism, this implies

$$\dim F_m' \cdot F_n' \geqq r - 1.$$

Another very interesting application of Theorem 3, or rather the corollary to Theorem 3, is the generalization of a theorem first proved by Miss Mullikin.† Miss Mullikin showed that the sum of a countable number of closed sets, no one of which separates the plane, and whose intersections taken pairwise are vacuous, cannot separate the plane. Recalling that the vacuous set is of dimension minus one, and the plane is R^2, we see that this is a particular case of the following theorem:

THEOREM 4. *In R^n let A_1', A_2', $\cdots$ be a countable collection of closed sets no one of which separates R^n, and such that*

$$\dim A_i' \cdot A_j' \leqq n - 3, \quad \textit{for} \quad i \neq j.$$

Under these conditions the sum $S=\sum_{i=1}^{\infty} A_i'$ cannot separate R^n.

Proof. Let us suppose, to the contrary, that S does separate R^n. Then

$$R^n - S = M_1 + M_2,$$

where M_1 and M_2 are non-vacuous and mutually separated. From the well known result that if a set X separates a connected set M then some closed subset Y of X also separates M,‡ it follows that S contains a closed subset A which separates M_1 from M_2 in R^n. Setting $A_i = A \cdot A_i'$, we have $A = \sum_{i=1}^{\infty} A_i$ (A_i closed). Taking a_1 a point of M_1 and a_2 a point of M_2, we have $a_1 - a_2 \cong 0$ in $R^n - A_i$ but $a_1 - a_2$ non-$\cong 0$ in $R^n - A$. It follows from the corollary to Theorem 3 that there exists a pair of integers m and n such that $\dim A_m \cdot A_n \geqq n - 2$. Since $A_i' \supset A_i$, $\dim A_m' \cdot A_n' \geqq n - 2$ contrary to the hypothesis of the theorem.

We return now to our discussion of the dimension of closed sets. If the

† A. M. Mullikin, *Certain theorems relating to plane connected point sets*, these Transactions, vol. 24 (1922), pp. 144–162.

‡ See Knaster and Kuratowski, Fundamenta Mathematicae, vol. 2 (1921), pp. 234–235.

dimension of the closed set F is r, then F is a simple r-dimensional obstruction at a point p of F. We can find an $\epsilon>0$ and a cycle z^{n-r-1} in $S(p, \epsilon)-F$ such that z^{n-r-1} does not bound in $S(p, \epsilon)-F$. Denote by d the distance between F and z^{n-r-1}. Now let F be decomposed into the sum of a countable number of closed sets $F_1, F_2, \cdots, F_s, \cdots$ each of which is of diameter smaller than d. These sets will satisfy the hypotheses of Theorem 3, if $z^{n-r-1}\cong 0$ in $S(p, \epsilon)-F_s$.

If $F_s\cdot S(p, \epsilon)=0$, the above is certainly true. Let us assume therefore that q_s is a point of $F_s\cdot S(p, \epsilon)$. $\delta_s=\delta(F_s)<d$. Choose two numbers α and β such that $\delta_s<\alpha<\beta<d$. $S(q_s, \alpha)\supset F_s$ and $S(p, \epsilon)-\overline{S(q_s, \alpha)}\supset z^{n-r-1}$. We may assume that $r\neq 0$, since the theorem we are about to prove is trivial for the case $r=0$. We now show that z^{n-r-1} can be deformed into a point in $S(p, \epsilon)-\overline{S(q_s, \alpha)}$, from which $z^{n-r-1}\cong 0$ in $S(p, \epsilon)-F_s$ follows. This may be done in two steps. First project z^{n-r-1} on $S(p, \epsilon)\cdot B(S(q_s, \beta))$, with center of projection q_s. Denote the projection of z^{n-r-1} by $z_{1/2}^{n-r-1}$. Since $S(p, \epsilon)\cdot B(S(q_s, \beta))$ is either equal to $B(S(q_s, \beta))$ or homeomorphic to a hemisphere of $B(S(q_s, \beta))$, our second step, deforming $z_{1/2}^{n-r-1}$ into a point is possible (note $n-r-1\neq n-1$). The projection can be considered a deformation with the parameter varying from 0 to $\frac{1}{2}$ as z^{n-r-1} moves to $z_{1/2}^{n-r-1}$ and from $\frac{1}{2}$ to 1 in the second step.

We can therefore say that if a set F is of dimension r, there exists a number d such that any decomposition of F into a countable infinity of closed sets of diameter less than d has the property that the intersection of at least one pair is of dimension at least $r-1$. Conversely, if a closed set F in R^n can be ϵ-decomposed into a countable infinity of closed sets with the intersection of any pair of dimension at most $r-1$, and if ϵ is arbitrary, then the dimension of F is at most r. From the fact that any closed set of dimension r can be decomposed into a countable infinity of closed sets of diameter less than any preassigned ϵ, whose intersections, taken pairwise are of dimension at most $r-1$ (if the set is compact, these may even be taken to be finite in number) we can state the following theorem:

THEOREM 5. *A closed subset F of R^n is of dimension r if and only if there exists an $\epsilon>0$, such that F may be decomposed into the sum of a countable infinity of closed sets $F_1, F_2, \cdots, F_s, \cdots$, of diameter less than ϵ with* $\dim F_i\cdot F_j\leqq r-1$, *$i\neq j$, but for any such decomposition there exists a pair of integers m and n such that* $\dim F_m\cdot F_n=r-1$.

UNIVERSITY OF PENNSYLVANIA,
PHILADELPHIA, PA.

CONCERNING LIMITING SETS IN ABSTRACT SPACES, II*

BY

R. G. LUBBEN

In his first paper on limiting sets† the author considered the distributive property in connection with metric spaces. In this paper we consider the property in connection with more general spaces and show that it and weak additional hypotheses imply that every uncountable point set in the space under consideration (1) is α-compact in itself‡ and (2) is separable. It is well known that in a metric space properties (1), (2), and the following, (3), are equivalent: (3) Every point set has the Lindelöf property. Sierpiński has shown that (2) and (3) are independent in a space S.§ In consideration of Sierpiński's result an equivalence involving these properties as stated in Theorem 7 is of considerable interest and is used in showing that (2) holds in Hausdorff space.

Above we discussed certain properties that hold "im grossen." With the help of the first countability axiom, or a more general hypothesis concerning

* Presented to the Society, December 27, 1928, August 30, 1929, and September 9, 1931; received by the editors September 8, 1936 and, in revised form, May 14, 1937.

† These Transactions, vol. 30 (1928), pp. 668–685. In a topological space the *limiting set* of an aggregate G of sets is the set of all points P of the space such that every neighborhood of P contains points in common with infinitely many distinct elements of G. The elements of G are understood to be sets $h(\alpha, g)$, where α is a number and g is a point set in the space; for the case $\alpha=0$ let $h(\alpha, g)$ be g, and for other values of α let the elements of $h(\alpha, g)$ be α and the points of g. Thus, we may refer to an element $h(\alpha, g)$ of G as a point set if $\alpha=0$. A topological space is said to have the *distributive property* provided that if in that space K is a closed point set, G is a collection of sets, and if each point of K belongs to some subset of K which is the limiting set of a sub-collection of G, then K itself is the limiting set of a sub-collection of G.

For well known definitions and for general information about topological spaces see: M. Fréchet, *Les Espaces Abstraits et leur Théorie Considérée comme Introduction à l'Analyse Générale*, 1928; F. Hausdorff, (I) *Grundzüge der Mengenlehre*, 1914, and (II) *Mengenlehre*, 1927; K. Menger, (I) *Dimensionstheorie*, 1928, and (II) *Kurventheorie*, 1932; R. L. Moore, *Foundations of Point Set Theory*, American Mathematical Society Colloquium Publications, vol. 13, 1932; C. Kuratowski, *Topologie*, 1933; W. Sierpiński, (I) *Introduction to General Topology*, 1934, translated by C. C. Krieger; and P. Alexandroff and H. Hopf, *Topologie*, 1935.

‡ A space or a point set is *α-compact* (*in itself*) provided that every uncountable point set in it has a limit point (contains a limit point of itself). Cf. W. Gross, *Zur Theorie der Mengen in denen ein Distanzbegriff definiert ist*, Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, part IIa, vol. 123 (1914), p. 805.

§ Cf. W. Sierpiński, (II) *Sur l'équivalence de trois propriétés des ensembles abstraits*, Fundamenta Mathematicae, vol. 2 (1921), pp. 179–188; C. Kuratowski and W. Sierpiński, *Le théorème de Borel-Lebesgue dans la théorie des ensembles abstraits*, Fundamenta Mathematicae, vol. 2 (1921), pp. 172–178.

monotonic families of neighborhoods* the distributive property gives also local compactness and regularity. The author has been unable to prove that the Lindelöf property is among these necessary conditions; if it is, it is possible to state simple necessary and sufficient conditions for the distributive property (see Theorems 14 and 16).

LEMMA I. *In a Fréchet space H in order that every point set be separable, it is necessary and sufficient that* (1) *every closed set be separable and* (2) *if a point is a limit point of a point set, it is a limit point of a countable subset of the point set.*

LEMMA II. *In a Fréchet space V every monotonic family of neighborhoods of a point contains a sub-collection which is a well-ordered monotonic descending family of neighborhoods of the point.*

THEOREM 1. *A space satisfies the first countability axiom if each point in it has a monotonic family of neighborhoods and one of the following holds*: (A) *The space is a Hausdorff space in which every point set is α-compact in itself*; (B) *the space is a space H in which a point is a limit point of a point set if and only if it is a limit point of a countable subset of the point set.*

Proof. Consider first case (A). Let S be the set of all points in the space, P be a point in it, H be a well-ordered monotonic descending family of neighborhoods of P, and K be a well-ordering of the points of $S-P$. Let U_1 be an element of H, P_1 be the first point of K in U_1, and V_1 the first element in H which is a subset of U_1 and of which P_1 is not a point or a boundary point. Suppose that P is not an isolated point of the space. Suppose that U_x, P_x, and V_x have been defined for each ordinal x less than a definite ordinal α. Provided that there exist elements of H common to all V_x's for $x<\alpha$ we shall define U_α, P_α, and V_α as follows: U_α is the first element of H common to all the V_x's for $x<\alpha$; P_α is the first element of K in U_α; V_α is the first element of H which is a subset of U_α and of which P_α is not a point or limit point. Let G be the well-ordered sequence $(V_1, V_2, V_3, \cdots, V_\omega, \cdots, V_\alpha, \cdots)$, where α is such that U_α, P_α, and V_α exist; and let $E=(P_1, P_2, P_3, \cdots, P_\omega, \cdots, P_\alpha, \cdots)$. Clearly each point of E is an isolated point of E. It follows from our condition that E, and hence G, has each a finite or a countable number of elements.

We shall show that each element of H contains an element of G. Suppose this were not true; then, since H is monotonic, it would contain an element U

* A *complete* family of neighborhoods of a point is one that defines the operation of derivation at that point; cf. Fréchet, pp. 172–173. Such a family is *monotonic* provided that every pair of its elements has the property that one is a subset of the other, and said to be *monotonic descending* with reference to a definite ordering provided that if one element precedes another, the first contains the second. In connection with a space H Fréchet we shall consider the term neighborhood as equivalent to the term "open set"; cf. Fréchet, pp. 186–187.

which is a subset of all elements of G. There exists a first ordinal λ which is greater than all ordinals x such that there exists an element V_x of G. Since H is a well-ordered sequence, there exists a first one of its elements that is common to all elements of G, and this element is by definition U_λ. Then the first point of K in U_λ is P_λ. There exist open sets R_1 and R_2 containing P_λ and P respectively and having no common points. Then P is not a point or limit point of $S-U_\lambda$ or of $S-R_2$; hence, there exists a first element of H that is a subset of U_λ and does not have P_λ on its boundary. This element is by definition V_λ. But, this is contary to the definition of λ. Thus, every element of H contains an element of G; the converse is true. Since H is a family of neighborhoods of P, so is G.†

Consider next case (B). Let P be a limit point of the set of distinct points, $E=(P_1, P_2, P_3, \cdots)$, none of which is P, and let $H=[W]$ be a family of neighborhoods of P. For each n let W_n be an element of H containing no point of $P_1+P_2+P_3+ \cdots +P_n$, and $G=(W_1, W_2, W_3, \cdots)$. Then each element of H contains an element of G; for, if some element U of H did not contain an element of G, it would be a subset of every element of G. Hence, if n is any integer, U is a subset of W_{n+1} and does not contain P_n. But this involves a contradiction, since P is a limit point of E. Thus every element of H contains an element of G, and conversely. It follows that G is a family of neighborhoods of P.†

THEOREM 2. *A locally compact Hausdorff space which has the Lindelöf property satisfies the first countability axiom.*

Proof. Let P be a point of our space T. For each point Q of $T-P$ let U_Q and V_Q be mutually exclusive open sets containing Q and P respectively. Then $T-P$ may be covered by a countable sequence $(U_{Q_1}, U_{Q_2}, U_{Q_3}, \cdots)$ of the elements of $[U_Q]$. Let R be an open set containing P such that $\overline{R}$ is compact; let $W_n=R\cdot(T-\sum_{i=1}^{i=n}\overline{U}_{Q_i})$; and let $F=(W_1, W_2, W_3, \cdots)$. Then F is a monotonic descending family of neighborhoods of P.

For, let M be a point set having points distinct from P in every element of F. It may be shown that M has a subset $N=(P_1, P_2, P_3, \cdots)$ of distinct points such that for $P_n \epsilon W_n$ each n. Since $R \supset W_n$, $\overline{W}_n$ is compact; hence N has a limit point X which is a point of $\overline{W}_1\cdot\overline{W}_2\cdot\overline{W}_3\cdot \cdots$. If X were a point of $T-P$, there would be an integer n such that X belongs to U_{Q_n}. Since $\overline{W}_n$ contains no point of U_{Q_n} we are involved in a contradiction.

Conversely if P is a limit point of a point set K, every element of F contains a point of K distinct from P.

LEMMA III. *In a space H the limiting set of a collection of point sets is closed.*

† Cf. Fréchet, p. 173.

THEOREM 3. *Every regular space H which satisfies the first countability axiom and has the distributive property is locally compact.*

This theorem may be proved by methods analogous to those used in the proof of Theorem 8 of the author's first paper, p. 677.

THEOREM 4. *Every space V which has the distributive property is α-compact.*

Proof. Suppose there exists a space which satisfies the hypothesis, but contains an uncountable point set M whose derived set is vacuous. Let K be a countable subset of M, $N=M-K$, and $P_1, P_2, P_3, \cdots$ be points of K. For each point x of N and each positive integer n let $g_{xn}=x+P_n$. Let G^* be the aggregate $[g_{xn}]$. For each point x of N there exists a neighborhood R_x of x which contains no point of $M-x$. Hence x is the limiting set of the aggregate $G_x^*=(g_{x1}, g_{x2}, g_{x3}, \cdots)$. Since N has no limit point, it is closed. It follows from our hypothesis that G^* contains a sub-collection G whose limiting set is N. Let $G_x=G\cdot G_x^*$. Since R_x contains no point of any element of $G-G_x$, G_x contains infinitely many distinct elements. Since N is an uncountable point set, and an element of G_y contains the point z of N only if $y=z$, G has uncountably many elements. Hence there exists an integer m such that P_m is common to infinitely many elements of G. This involves a contradiction with the fact that N is the limiting set of G.

THEOREM 4A. *In order that a metric space should have the distributive property, it is necessary and sufficient that it be locally compact and separable.*

This theorem is a consequence of Theorem 4, Gross, loc. cit., pp. 805–806, and Theorems 8 and 9 of the authors first paper, pp. 677–678.

A space S_1 is said to be a *sub-space* of a space S_2 provided that (1) every point of S_1 is a point of S_2, and (2) if P is an arbitrary point and M an arbitrary point set in S_1 then P is a limit point of M in S_1 if and only if it is a limit point of M in S_2.

LEMMA IV. *Every subspace of a space S is a space S, and every subspace of a space H is a space H.*

THEOREM 5. *If a space S has the distributive property, then every regular, locally compact subspace of it has this property.*

Proof. Let S_1 be a space S having the distributive property, and let T be a regular, locally compact sub-space of it. In T let K be a closed point set and G be a collection of sets such that each point P of K belongs to a subset K_P of K which is the limiting set of a sub-collction G_P of G. Let L_P be the limiting set of G_P with respect to the space S_1. Then $K_P=L_P\cdot M$, where M is the set of all points belonging to T. Let N be the sum of all point sets L_P, where

the range of P is K; N' be the derived set of N with reference to the space S_1; and $\overline{N}=N+N'$. It follows from the definition of N that $N\cdot M=K$. Suppose that $N'\cdot M$ contains a point Q which does not belong to K. Since K is a closed point set with respect to T, Q is not a limit point of K in either T or S_1. Hence, there exists in T an open set R_1 containing Q such that if $\overline{R}_{1T}$ denotes the sum of R_1 and its limit points in T, then $K\cdot\overline{R}_{1T}$ is vacuous. Since T is locally compact, there exists in it an open set R_2 containing Q such that $\overline{R}_{2T}$ is compact. Let $R_3=R_1\cdot R_2$. Then R_3 is an open set in T and contains Q. Also $\overline{R}_{1T}\supset\overline{R}_{3T}$. It follows that Q is not a limit point of $M-R_3$ in the space S_1. Hence, there exists in S_1 an open set U which contains Q but contains no point of $M-R_3$. Since Q is a limit point of N, there exists in K a point x such that U contains a point y of L_x. Then U must contain points of infinitely many elements of G_x. Since the elements of G_x are subsets of M, and $U\cdot M$ is a subset of R_3, R_3 contains points of infinitely many elements of G_x. Since $\overline{R}_{3T}$ is compact, it must contain a point W which belongs to the limiting set of G_x; since $\overline{R}_{1T}\supset\overline{R}_{3T}$, the latter contains no point of K. This involves a contradiction with the fact that $K\supset K_x$. Hence the point Q does not exist, and $\overline{N}\cdot M=K$.

For each point x of $\overline{N}-K$ let h_{1x}, h_{2x}, h_{3x}, $\cdots$ be the pairs $(x, 1)$, $(x, 2)$, $(x, 3)$, $\cdots$. Let H be the aggregate $[h_{ix}]$, where the range of i is the set of positive integers, and that of x is the point set $\overline{N}-K$. Since S_1 is a space S, $\overline{N}$ is closed, and for every point P of K the following holds: $\overline{N}\supset N\supset L_P$. It follows that each point of $\overline{N}$ belongs to a subset of $\overline{N}$ which is the limiting set in S_1 of a sub-collection of $G+H$. Since S_1 has the distributive property, $G+H$ has a sub-collection G_1+H_1 such that $\overline{N}$ is the limiting set of G_1+H_1 in S_1 and such that G_1 and H_1 respectively are sub-collections of G and H respectively.

Suppose that K is not a subset of K_1, where K_1 is the limiting set of G_1 in S_1. Then K must contain a point E which belongs to the limiting set of H_1. Let R_E be an open set in T containing E such that $\overline{R}_{ET}$ is compact. It follows from an analogous situation above that in S_1 the point E is not a limit point of $M-R_E$ and that there exists in S_1 an open set U_E which contains no point of $M-R_E$. Then U_E contains a point X of an element of H_1. Let B denote the limiting set of G in S_1. Since $B\supset N$, and B is closed, $B\supset\overline{N}$. Since X is a point of $\overline{N}-K=\overline{N}-\overline{N}\cdot M$, and all elements of G are sub-sets of M, X is the unique limit point in S_1 of an infinite subset of M, and U_E, which contains X, contains such a set E_X. Then $R_E\supset M\cdot U_E\supset E_X$. Since X is not a point of T, E_X has no limit point in T. This involves a contradiction with the fact that $\overline{R}_{ET}$ is a compact subset of T. Hence $K_1\supset K$. Since $M\supset K$, it follows that $M\cdot K_1\supset K$.

Since $\overline{N}\supset K_1$, $K=\overline{N}\cdot M\supset K_1\cdot M$. From $K\supset K_1\cdot M$ and $K_1\cdot M\supset K$, it fol-

lows that $K=K_1\cdot M$. Hence K is the limiting set of G_1 in T. Thus we have shown that T has the distributive property.

A space H is said to be *nearly a space* L provided that if in that space P is any point, M is any point set, and P is a limit point of M, then P is the derived set of a subset of M.

THEOREM 5A. *Every regular, locally compact subspace of a space H which has the distributive property and is nearly a space L has the distributive property.*

The proof is the same as that for Theorem 5.

THEOREM 6. *In a space H which has the distributive property every point set is α-compact in itself.*

Proof. Suppose the theorem is not true and that S_1 is a space H which has the distributive property but contains an uncountable point set M which contains no limit point of itself. Let T be the subspace of S_1 whose points are the points of M. To show that T has the distributive property adopt the notation of the proof of Theorem 5 and follow this proof to the place where the existence of the collections G_1 and H_1 is established, and suppose as there that K contains a point E not belonging to the limiting set of G_1. Define M' as the derived set of M in S_1. Since the points of elements of H_1 are points of $\overline{N}-K$, it follows that E is a limit point of $M'-M\cdot M'$ in S_1. Since S_1 is a space H, derived sets in it are closed, and E is a point of M' and hence a limit point of M. Since E is a limit point of M, we are involved in a contradiction. Thus E does not exist, and the argument of Theorem 5 shows that T has the distributive property.

By Theorem 4 the set M must have a limit point in T. But this again is contrary to the definition of M. Thus the supposition that the theorem is not true leads to a contradiction.

Note. When the space of our hypothesis is a space S the proof may be simplified. Let S_1 be our space and define M and T as above; then T is a regular, locally compact subspace of S_1, since all its points are isolated. By Theorem 5, T has the distributive property, and by Theorem 4, we are involved in a contradiction.

THEOREM 7. *In order that in a space H each point set either be condensed in itself or be separable, it is necessary and sufficient that every point set be α-compact in itself.*†

Proof. Obviously the condition is necessary. Suppose that it is not sufficient, and that the space contains a point set E which neither contains a con-

† In part our proof of Theorem 7 follows methods used by Sierpiński; cf. Sierpiński (II). See also the introduction for a discussion of the relation of Theorem 7 to some results by Sierpiński and Kuratowski.

densation point of itself nor is separable. Then, for each point P of E there exists a countable subset $D(P)$ of E such that P is not a point or a limit point of $E-D(P)$. Let T be a well-ordered sequence $(p_1, p_2, p_3, \cdots, p_\omega, p_{\omega+1}, \cdots, p_\alpha, \cdots)$ of the points of E. We shall now define a sequence of the type δ, where δ is the smallest transfinite ordinal of the third class; $U=(q_1, q_2, q_3, \cdots q_\omega, q_{\omega+1}, \cdots, q_\beta, \cdots)$, where $\beta<\delta$.

Proceed as follows: Let $q_1=p_1$. Let β be a definite ordinal less than δ. Suppose that q_x has been defined for all ordinals $x<\beta$, and let U_β be the set of all q_x's for such x's. Let S_β be the sum of all point sets $D(q_y)$, where q_y is an element of U_β. Let q_β be the first point of T which is not a point or a limit point of S_β.

We shall now show that q_β exists for every ordinal β less than δ. For, if q_β does not exist for all such ordinals β, there must be a first such ordinal λ for which it does not exist. Then it follows from our definitions that each point of E is either a point or a limit point of S_λ. But S_λ is the sum of all point sets $D(q_z)$, where q_z ranges over U_λ; thus S_λ is the sum of a countable number of countable sets and is countable; then E is separable. Thus, the supposition that λ exists leads to a contradiction.

Next we shall show that q_β is an isolated point of U. By definition q_β is not a point or a limit point of U_β. Further, $S_{\beta+1}$, which contains $D(q_\beta)$, contains no point of $U-U_{\beta+1}$. Since $U-q_\beta=U_\beta+(U-U_{\beta+1})$, it follows that q_β is an isolated point of U.

Thus, every point of the uncountable sequence U is an isolated point of U. By our hypothesis, however, U must contain a limit point of itself. Thus, the supposition that our condition is not sufficient has led to a contradiction.

THEOREM 8. *In order that for each infinite collection of point sets in a space H it be true that at most a countable number of its elements fail to be subsets of its limiting set, it is necessary and sufficient that every point set in the space be α-compact in itself.*

Proof. We shall first show that the condition is sufficient. Suppose that it is not and that there exists in our space a point set K and a collection G of point sets such that K is the limiting set of G, but that G contains an uncountable sub-collection G_1, none of whose elements are subsets of K. For each element g of G_1 let P_g be a point which does not belong to K. It follows by our hypothesis that the set $[P_g]$ *contains* a point Q, every neighborhood of which contains infinitely many elements of $[P_g]$. Then Q belongs to the limiting set of G, and we are involved in a contradiction.

Conversely, let M be an uncountable point set in a space which satisfies our condition. Let G_2 be a collection of point sets whose elements are the

points of M, no two elements being the same point. Then M' is the limiting set of G, and M' and M have uncountably many points in common. Thus, the condition is necessary.

Theorems 8 and 9 are generalizations of Theorems 2 and 4, respectively, of our first paper, and are of interest in connection with Theorem 7, and also with Theorems 6 and 10A, in that they indicate consequences of the distributive property.

THEOREM 9. *In order that for a space H every infinite collection of sets should contain a countable sub-collection having the same limiting set as the collection itself, it is necessary and sufficient that every point set in the space be separable.*

Proof. To prove the necessity of the condition proceed as follows: Let N be a point set and $M=\overline{N}$. Let G be a collection such that for each point x of N there exists a collection of elements of G, $g_{1x}, g_{2x}, g_{3x}, \cdots$, where g_{nx} is the pair (n, x). Now proceed by methods analogous to those used in the proof of Theorem 4 of the author's first paper.

Consider next the sufficiency. Let G be a collection of sets and K be the limiting set of G. By our condition K contains a countable subset $N=P_1+P_2+P_3+\cdots$ such that $\overline{N}=K$. Then *for each n the point P_n belongs to the limiting set of some countable sub-collection of G.* Suppose that for some definite $P_j=Q$ this is not true. Let g_0 be a definite element of G. For n greater than zero suppose that g_k has been defined for $k<n$. Let $G_n=G-(g_0+g_1+g_2+\cdots+g_{n-1})$; let H_n be the sum of all elements of G_n; and let $F_n=Q_1+Q_2+Q_3+\cdots$ be a countable set such that $\overline{F}_n\supset H_n\supset F_n$. Then Q is a point of F_n'.

For, let R be an open set containing Q. Since Q belongs to the limiting set of G, and hence of G_n, R contains points of infinitely many elements of G_n. If Q were common to infinitely many elements of G_n, it would be common to a countable infinity of such elements and would belong to the limiting set of this countable collection; this, however, is contrary to the definition of Q. Thus R contains points of H_n distinct from Q, and hence points of F_n. Thus $Q\epsilon F_n'$. For each positive integer k let t_k denote a definite element of G_n that contains Q_k. Let T denote the aggregate $(t_1, t_2, t_3, \cdots)$. Since T has a finite or a countable number of elements, Q does not belong to its limiting set. Hence, there must exist an open set U containing Q which contains points of at most a finite number of the elements of T, say of $t_{k_1}, t_{k_2}, \cdots, t_{k_j}$. Then $U\cdot F_n=\sum_{i=1}^{i=j}U\cdot F_n\cdot t_{k_i}$. Since Q belongs to the derived set of $U\cdot F_n$, it follows that for some i, $0<i<j+1$, Q is a limit point of $U\cdot F_n\cdot t_{k_i}$, that is of t_{k_i}. Let g_n be defined as the sum of n and such a t_{k_i}, and $E=(g_1, g_2, g_3, \cdots)$. Then Q is a limit point of each element of E, and hence belongs to the limiting set of E.

Thus, each point P_n of N is a point of the limiting set of some countable sub-collection M_n of G. Let $M=\sum_{t=1}^{t=\infty} M_n$, and let L be the limiting set of M. Then $K=\overline{N}\supset L\supset N$. Since in a space H derived sets are closed, it follows that $L=\overline{N}=K$. But, M is a countable sub-collection of G.

THEOREM 10. *In a Hausdorff space having the distributive property every closed point set is separable.*

Proof. Suppose that a space S_1 satisfies the hypothesis but not the conclusion of our theorem. Then there exists in it an uncountable, closed, non-separable point set E. By Theorems 6 and 7 the set E contains a point of condensation of itself Q; similarly $E-Q$ contains a point of condensation of itself P. Let U and V be mutually exclusive open sets containing Q and P, respectively. Let $M=U\cdot E$, $N=V\cdot E$, and $H=[S_1-(U+V)]\cdot E$. Then one of the three point sets M, N, or H is non-separable; for otherwise E, their sum, would be separable. Consider the two cases: (I) Either M or N is not separable; (II) both M and N are separable. Consider first case (I) and suppose that it is N that is not separable. Let $K=E-\overline{M}$; then $H+N=E-M\supset K\supset N$, and K is not separable. For, suppose that K is separable and has a countable subset K_1 such that $\overline{K}_1\supset K$. By definition H is the product of the two closed sets E and $S_1-(U+V)$, and thus is closed. It follows that every point of $N=K-K\cdot H$ is either a point or a limit point of $N\cdot K_1$. Thus, the supposition that K is separable, involves a contradiction with the hypothesis that N is not separable.

We shall now define certain sequences by an induction process. Let z_0 be a point of $M-Q$, $R_0=z_0$, and $U_0=U$. Now suppose that U_k, R_k, and z_k have been defined for all non-negative integers k less than the definite integer n. Let U_n be an open set containing Q such that $U_{n-1}-\overline{R}_{n-1}\cdot U_{n-1}\supset U_n$, let z_n be a point of $M\cdot(U_n-Q)$, and let R_n be an open set containing z_n such that $\overline{U}_n-Q\supset\overline{R}_n$. Let $F=z_1+z_2+z_3+\cdots$. The existence of U_n, z_n, and R_n for every positive integer n may be shown by making use, in particular, of Hausdorff's Axiom D. Since the open sets $R_1, R_2, R_3, \cdots$ are mutually exclusive, it follows that $F\cdot F'$ is vacuous.

For each point t of $\overline{K}$ and each positive integer n let $g_{tn}=t+z_n$, and G be the aggregate of all such g_{tn}'s. Since $z_1, z_2, z_3, \cdots$ are distinct points, the limiting set of the aggregate $(g_{t1}, g_{t2}, g_{t3}, \cdots)$ is $t+F'$. Thus, each point of $\overline{K}+F'$ belongs to a subset of $\overline{K}+F'$ which is the limiting set of a sub-collection of G. Hence, G has a sub-collection G_1 whose limiting set is $\overline{K}+F'$. Let W be the product of $\overline{K}$ and the sum of the elements of G_1. Since $E-\overline{M}=K$, and $M\supset F$, it follows that $K\cdot\overline{F}$ is vacuous. Then every point of K is a point or a limit point of W, and so is every point of $\overline{K}$. If the elements of G_1 were count-

able, so would be the points of W; then $\overline{K}$ would be separable. This is impossible, since K is not separable, and no point of K is a limit point of $\overline{K}\ \ K$. Hence, there are uncountably many elements of G_1, and there must exist an integer j such that z_j belongs to uncountably many elements of G_1. Then z_j belongs to the limiting set of G_1, that is to $\overline{K}+F'$. But, z_j belongs to neither $\overline{K}$ nor F'. Thus, case (I) involves a contradiction.

Consider next case (II). Since both M and N are separable, so are $\overline{M}$ and $\overline{N}$. Let $K=E-(\overline{M}+\overline{N})$. Then K is not separable. Define first F and then G precisely as in the proof of case (I); by following this proof we again arrive at a contradiction. Thus, the supposition that the theorem is not true is untenable.

THEOREM 10A. *If a Hausdorff space is nearly a space L and has the distributive property, every point set in it is separable.*

This is a consequence of Theorems 10 and 6 and Lemma I.

THEOREM 11. *A space H which satisfies the first countability axiom is a Hausdorff space if and only if it is a space S.*†

THEOREM 12. *A locally compact space S (Hausdorff space) which satisfies the first countability axiom is regular.*‡

THEOREM 12A. *If a Hausdorff space is locally compact at one of its points P, and P has a countable family of neighborhoods, the space is regular at P.*

THEOREM 13. *A space S (Hausdorff space) which has the distributive property and satisfies the first countability axiom is regular.*

Proof. Suppose that S_1 is a space S which satisfies the hypothesis of the theorem but contains a point P at which it is not regular. Then there exists in S_1 an open set R containing P such that if R_i is any open set whatever containing P, then $\overline{R}_i$ is not a subset of R. Let $(V_1, V_2, V_3, \cdots)$ be a countable family of neighborhoods containing P such that $R \supset V_1 \supset V_2 \supset V_3 \supset \cdots$. Let $U_1=V_1$, $m_1=1$, and M_1 be a countable subset of V_1-P which has a unique limit point in $R'-R$, say P_1. By an induction process we shall now define U_n, M_n, m_n, and P_n for every positive integer n. Proceed as follows: Suppose they have been defined for all n's less than a definite integer k. Let m_k be the first integer greater than m_{k-1} such that V_{m_k} contains no point of $\sum_{j=1}^{j=k-1} M_j$, let $U_k=V_{m_k}$, and let M_k be a compact countable sequence of points belonging to U_k-P and having a unique limit point P_k belonging to $R'-R$.

† Cf. Hausdorff (I), pp. 263–265, and Fréchet, *Démonstration de quelques propriétés des ensembles abstraits*, American Journal of Mathematics, vol. 50 (1928), p. 65.

‡ Cf. Alexandroff and Urysohn, *Mémoire sur les espaces topologique compacts*, Verhandelingen, Koninklyjke Akademie van Wetenschappen, Amsterdam, vol. 14 (1929), pp. 28–29.

We shall now prove that U_k, M_k, m_k, and P_k exist for all positive integers k; the argument suggests, in particular, the proof for the case $k=1$. Suppose that our proposition has been established for all n's less than k, where $1<k$. Then each of the sequences $M_1, M_2, M_3, \cdots, M_{k-1}$ has exactly one limit point; the derived set of their sum is $P_1+P_2+\cdots+P_{k-1}$, which is a subset of $R'-R$. Then P is not a limit point of the closed point set $\sum_{j=1}^{j=k-1}M_j$, and there exists an integer m_k greater than m_{k-1} such that V_{m_k} contains no point of this point set. Hence U_k exists. It follows from the definition of R that $\bar{U}_k$ is not a subset of R, and that U_k has a limit point P_k in $R'-R$. Since our space is a space S, U_k-P has a countable subset M_k such that P_k is the unique limit point of every infinite subset of M_k. Our existence theorem may thus be established by mathematical induction.

The sequence $P_1, P_2, P_3, \cdots$ contains a sub-sequence $P_{n_1}, P_{n_2}, P_{n_3}, \cdots$, having not more than one limit point, such that $n_1<n_2<n_3<\cdots$. If this sub-sequence has a limiting set, let it be denoted by the symbol Q; otherwise, let Q be the null set. Let $P_{n_k}=Q_k$; $M_{n_k}=N_k$; let $O_{1k}, O_{2k}, O_{3k}, \cdots$ be the points of N_k; $g_{jk}=Q+O_{k1}+O_{jk}$; let G_k be the sequence $(g_{1k}, g_{2k}, g_{3k}, \cdots)$; $G^*=\sum_{j=1}^{j=\infty}G_j$; and $K=Q+N_1+\sum_{k=1}^{k=\infty}Q_k$. The limiting set of G_k is $Q+O_{k1}+Q_k$. Since K is closed and the space has the distributive property, G^* contains a sub-collection G whose limiting set is K. Suppose that G contains elements in common with at most a finite number of the elements of the aggregates $G_1, G_2, G_3, \cdots$, say with those having subscripts not greater than a definite integer t. Then, contrary to the fact that K contains infinitely many distinct points, the limiting set of G is a subset of $Q+\sum_{k=1}^{k=t}(O_{k1}+Q_k)$. Hence, for infinitely many values of k there exist elements of G which contain points of N_k. Thus, every element of $(V_1, V_2, V_3, \cdots)$ contains points in common with infinitely many distinct elements of G, and P belongs to the limiting set of G. Thus, the supposition that our space is not regular has led to a contradiction.

THEOREM 13A. *If a space S has the distributive property and a point P in it has a countable family of neighborhoods, the space is regular at P.*

This theorem may be proved by the methods used for Theorem 13. It follows by Theorem 3 that the space is locally compact at P.

Note. The statements of Theorems 12 and 13 differ only in that in the hypothesis of the one the distributive property takes the place of local compactness in that of the other; they are stated for both spaces S and Hausdorff spaces. Theorems 12A and 13A are analogous generalizations of Theorems 12 and 13 respectively; but 12A is stated for a Hausdorff space, while 13A is stated for a space S. The question arises as to whether each of the Theorems 12A and 13A hold for both types of spaces. The author has not found the

answer for the case of 13A; for 12A the answer is in the negative, as may be seen by considering a space T, when points and limit points are those of $P+K+\sum_{i=1}^{i=\infty} N_i$ of the proof of Theorem 13.

THEOREM 14. *If a Hausdorff space or a space S has the distributive property and each point in it has a monotonic family of neighborhoods, then the following properties hold for the space*: (1) *It satisfies the first countability axiom*; (2) *it is both a Hausdorff space and a space S*; (3) *every point set in it is separable*; *and* (4) *it is regular and locally compact.*

Proof. By Theorems 6 and 1 our space satisfies the first countability axiom; by Theorem 11 it is both a space S and a Hausdorff space; by Theorem 10 and Lemma I every point set in it is separable; by Theorem 13 it is regular; and by Theorem 3 it is locally compact.

Note. Theorems 14 and 15 may be regarded as a summary of results of this paper with regard to conditions necessary for the distributive property. Theorem 16 deals with sufficient conditions; Theorems 15 and 17 with necessary and sufficient conditions.

THEOREM 15. *If a space S (Hausdorff space) satisfies the first countability axiom and has the distributive property, then in order that one of its sub-spaces have the distributive property it is necessary and sufficient that the sub-space be regular and locally compact.*

Proof. The necessity follows from Theorem 14; the sufficiency from Theorem 5.

THEOREM 16. *A sufficient condition that a Hausdorff space have the distributive property is that it be locally compact and have the Lindelöf property, and that every closed point set in it be separable.*

Proof. By Theorem 2 our space satisfies the first countability axiom; by Theorem 12 it is regular; by Theorem 11 and Lemma I every point set in it is separable. The proof may be completed by following the methods for Theorem 9 of the author's first paper, p. 678. Theorem 17 follows from Theorems 14 and 16.

THEOREM 17. *In order that a Hausdorff space which has the Lindelöf property and in which every point has a monotonic family of neighborhoods should have the distributive property, it is necessary and sufficient that the space be locally compact and that every closed set in it be separable.*

UNIVERSITY OF TEXAS,
AUSTIN, TEXAS

A CORRECTION TO THE PAPER "ON EFFECTIVE SETS OF POINTS IN RELATION TO INTEGRAL FUNCTIONS"*

BY

V. GANAPATHY IYER

An additional hypothesis is necessary for the truth of Lemma 3. In the relation $g'(z_n) = P_n'(z_n)Q_n(z_n)$, in order that Lemma 3 may be true it is necessary to prove that $\lim_{n\to\infty}(\log|P_n'(z_n)|/|z_n|^\rho)=0$. Under the conditions stated in the Lemma we can prove only that $\limsup_{n\to\infty}(\log|P_n'(z_n)|/|z_n|^\rho)\leqq 0$. If we assume also that $\liminf_{n\to\infty}(\log\mu_n/|z_n|^\rho)\geqq 0$, where $\mu_n=\min_{n\neq\nu}|z_n-z_\nu|$, it will follow that $\liminf_{n\to\infty}(\log|P_n'(z_n)|/|z_n|^\rho)\geqq 0$, hence Lemma 3 will hold. This additional condition is obviously satisfied in the particular case worked out in Lemma 4 since the circles are non-overlapping after a certain stage and each circle contains only one zero. If μ_n is sufficiently small, Lemma 3 will cease to be true, hence one of the doubtful points raised in §3.6 is answered in the negative.

---

* Received by the editors December 6, 1937. Cf. these Transactions, vol. 42 (1937), pp. 358–365.

MADRAS UNIVERSITY,
MADRAS, S. INDIA